Theoretical Computer Science 15 (1981) 77–110
North-Holland Publishing Company

RESOLUTION DES SYSTEMES D'EQUATIONS ALGEBRIQUES

Daniel LAZARD
Departement de Mathématiques, Université de Poitiers, 86022 Poitiers Cedex, France

Communicated by M. Nivat
Received October 1979
Revised March 1980

Abstract. Let $f_1, \ldots, f_k$ be k multivariate polynomials which have a finite number of common zeros in the algebraic closure of the ground field, counting the common zeros at infinity. An algorithm is given and proved which reduces the computations of these zeros to the resolution of a single univariate equation which degree is the number of common zeros. This algorithm gives the whole algebraic and geometric structure of the set of zeros (multiplicities, conjugate zeros, ...). When all the polynomials have the same degree, the complexity of this algorithm is polynomial relatively to the generic number of solutions.

1. Introduction

La résolution des systèmes d'équations algébriques est un problème crucial de l'algorithmique algébrique. Nous n'envisagerons pas ici les methodes issues de l'analyse (méthode de Newton, optimisation, etc...) qui sont rapides, mais ne donnent aucun renseignement sur le nombre, où même l'existence des solutions.

Les méthodes algébriques, au contraire, peuvent donner tous les renseignements désirables sur l'espace des solutions (dimension, degré, multiplicité, corps de définition, etc...) mais sont extrêmement lentes. Aussi, pour améliorer leurs performances, a-t-on été amené à utiliser tout l'arsenal des techniques de manipulation symbolique (cf. [12]).

Les techniques antérieurement développées pour la résolution des systèmes d'équations algébriques se ramènent toutes à éliminer les inconnues l'une après l'autre à l'aide de résultants ou de méthodes analogues. Cela amène à manipuler des polynômes de degré très élevé afin de calculer des résultants de résultants. Il en résulte que la complexité du problème, qui est déjà très grande, est augmentée dans une proportion considérable; nous reviendrons plus bas sur cette question.

Il est par ailleurs surprenant de constater que les techniques modernes de géométrie algébrique ne sont pas utilisées pour résoudre explicitement les systèmes d'équations algébriques; c'est très étonnant, car la géométrie algébrique est généralement définie comme l'étude de l'ensemble des solutions de tels systèmes.

0304-3975/81/0000-0000/$02.50

La méthode de résolution qui est développée dans cet article est nouvelle; elle tire ses origines, d'une part de la 'théorie classique de l'élimination', et en particulier de la méthode du U-résultant [11, 8], et d'autre part des travaux de l'auteur sur la résolution des systèmes d'équations linéaires sur les anneaux de polynômes [6]. Cette méthode est 'géométrique', en ce sens qu'elle ne détruit pas la nature géométrique du problème; en particulier, l'invariance par transformation linéaire sur les inconnues est conservée.

Cette méthode consiste essentiellement à écrire une certaine matrice rectangulaire de polynômes du premier degré et à la réduire selon une méthode voisine de celle de Gauss afin d'obtenir une matrice carrée dont les coefficients sont encore du premier degré. Il se trouve que le déterminant de cette matrice est produit de facteurs linéaires qui sont en correspondance avec les solutions, les coordonnées d'une solution étant les coefficients du facteur correspondant.

Cet algorithme ne fonctionne que quand l'espace des solutions est de dimension zéro; comme de tels espaces sont 'triviaux' en géométrie algébrique, les techniques géométriques employées sont relativement élémentaires, et se limitent aux relations entre espaces affine et projectif, à la notion de multiplicité et au théorème des zéros de Hilbert. Les techniques algébriques se ramènent essentiellement aux propriétés des anneaux et modules gradués de dimension un. Il faut y ajouter le théorème permettant de borner les degrés à considérer (théorème 3.3) dont la démonstration utilise l'homologie du complexe de Koszul.

La rédaction de cet article posait quelque problème, car il n'était pas possible de séparer la description de l'algorithme, sous une forme facilement programmable, de son interprétation dans un langage abstrait qui, seule, permet de justifier les opérations effectuées. Nous avons essayé de résoudre cette difficulté en structurant l'article de la manière suivante:

Le paragraphe 2 décrit comment l'algorithme opère sur un exemple particulièrement simple.

Au paragraphe 3, nous posons le problème d'une manière précise, et le transcrivons dans le langage algébrique qui permet de justifier l'algorithme. Aux paragraphes 4 et 5, nous décrivons et justifions l'algorithme; cette description étant assez abstraite, nous en donnons une forme aisément programmable dans l'appendice A.4.

Au paragraphe 6, nous expliquons comment terminer les calculs, notamment en utilisant à nouveau l'algorithme du paragraphe 5. Le paragraphe 7 est consacré à l'étude de la multiplicité des solutions.

L'objet du paragraphe 8 est l'étude de la complexité de notre algorithme et de sa comparaison avec celle de la méthode classique. En conclusion (paragraphe 9), nous discutons les avantages de notre méthode.

Enfin, nous regroupons dans les appendices A.1, A.2 et A.3 les démonstrations des théorèmes du paragraphe 3; bien que les deux premiers de ces théorèmes soient, en principe, 'bien connus', nous n'avons pas trouvé les énoncés dont nous avions besoin dans la littérature. Le troisième est implicitement démontré dans [6], mais il nous a semblé utile d'en donner une démonstration directe.

2. Un exemple

Dans ce paragraphe, nous montrons comment l'algorithme qui fait l'objet de cet article opère sur un exemple particulièrement simple. Nous ne donnons guère de justifications et nous nous référons souvent à la suite de l'article. Cependant, nous pensons que cet exemple peut aider à la compréhension des paragraphes ultérieurs.

Considérons le système, très simple, d'équations

$$x^2+xy+2x+y-1=0, \tag{1}$$

$$x^2-y^2+3x+2y-1=0 \tag{2}$$

qui a trois solutions à distance finie, (0, 1), (1, −1) et (−3, 1) et une solution à l'infini dans la direction (−1, 1) (direction asymptotique commune).

Désignons par f_0 et f_1 les premiers membres de (1) et (2); ainsi

$$f_0=-1+2x+y+x^2+xy, \qquad f_1=-1+3x+2y+x^2-y^2.$$

Posons

$$f_2=U+Vx+Wy$$

où U, V et W sont des indéterminées.

L'entier D introduit par le théorème 3.3 étant égal à trois, contruisons le tableau suivant où les points représentent des zéros:

	1	x	y	1	$\underline{x}$	y	1	x	y	x^2	xy	y^2
1	−1	.	.	−1	.	.	U	.	.	.	.	.
x	2	−1	.	3	−1	.	V	U	.	.	.	.
y	1	.	−1	2	.	−1	W	.	U	.	.	.
x^2	1	2	.	1	3	.	.	V	.	U	.	.
$\underline{xy}$	1	1	2	.	2	3	.	W	V	.	U	.
y^2	.	.	1	−1	.	2	.	.	W	.	.	U
x^3	.	1	.	.	1	.	.	.	.	V	.	.
x^2y	.	1	1	.	.	1	.	.	.	W	V	.
xy^2	.	.	1	.	−1	.	.	.	.	.	W	V
y^3	.	.	.	.	.	−1	.	.	.	.	.	W

(3)

Voici comment ce tableau est construit: la colonne extérieure est constituée par les monômes de degré inférieur ou égal à $D=3$. Les trois parties du tableau correspondent aux polynômes f_0, f_1 et f_2. La ligne extérieure de la partie correspondant à f_i est constituée par les monômes de degré 3-degré (f_i). Si m est un monôme de la colonne extérieure et n un monôme de la ligne extérieure de la partie correspondant à f_i, alors le coefficient qui apparaît à l'intersection de la ligne et de la colonne correspondantes est le coefficient de m dans nf_i; ainsi, 2 est le coefficient de xy dans xf_1.

En effectuant une élimination de Gauss sur la matrice qui apparaît sur la partie gauche du tableau, celui-ci devient

$$\begin{array}{cccccccc}
1 & . & . & . & . & . & . & . \\
. & 1 & . & . & . & . & . & . \\
. & . & 1 & . & . & . & . & . \\
. & . & . & 1 & . & . & . & . \\
. & . & . & . & 1 & . & . & . \\
. & . & . & . & . & 1 & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & .
\end{array}
\begin{array}{cccccc}
. & . & . & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
U+V-W & U+V-W & -U-V+W & U & -U & U \\
-V+W & -U+V & U & U-4V+W & V & 0 \\
-2V+3W & -2U+W & 3U+V & -3V & U-W & V \\
-V+2W & -U-V+W & 2U+V & -U & U & W
\end{array} \tag{4}$$

Après une réduction de Gauss sur la matrice des coefficients de U dans (4), la partie non triviale de (4) devient

$$\begin{array}{cccccc}
U+V+W & 0 & 0 & -V+W & V-W & -V+W \\
-V+W & -U+2V-W & V & V-W & -V+W & V-W \\
W & -2V & U-V+W & V-W & -V+W & V-W \\
0 & 0 & 0 & -V+W & V-W & -V+W
\end{array} \tag{5}$$

Des transformations évidentes sur les lignes et les colonnes font apparaître des zéros supplémentaires:

$$\begin{array}{cccccc}
U+V-W & 0 & 0 & 0 & 0 & 0 \\
-V+W & -U+2V-W & V & 0 & 0 & 0 \\
W & -2V & U-V+W & 0 & 0 & 0 \\
0 & 0 & 0 & -V+W & 0 & 0
\end{array} \tag{6}$$

Le déterminant de la partie gauche de (6) est

$$(U+V-W)(V-W)(U-3V+W)(U+W). \tag{7}$$

Les coefficients des facteurs de (7) donnent les coordonnées projectives des solutions, la coordonnée 'à l'infini' étant le coefficient de U, le terme constant de f_2. Ces solutions sont donc:

$$(1, 1, -1), \quad (0, 1, -1), \quad (1, -3, 1), \quad (1, 0, 1),$$

la seconde étant la solution à l'infini.

Remarques. (i) Plutôt que de considérer globalement le tableau (3), comme nous venons de le faire, l'algorithme que nous décrivons ci-dessous procède en deux étapes: d'abord réduction de la partie numérique et calcul de la matrice C de la réduction, puis calcul des dernières lignes (ici 4) du produit de C et de la partie droite. Pour économiser de la place en mémoire, il y a lieu de calculer cette matrice C en écrivant successivement dans la même zone mémoire les différentes colonnes de la partie gauche de (3); c'est ainsi que nous avons procédé dans l'appendice A4.

(ii) Nous avons interprété l'intérieur du tableau (3) comme une matrice, dont les lignes et les colonnes sont indexées par des monômes (ligne et colonne extérieures). On retrouve donc la définition générale des matrices de [1]. Il est, certes, facile de numéroter les monômes à l'aide d'un ordre lexicographique (nous l'avons fait implicitement pour représenter la matrice), mais, ce faisant, on perd de l'information; en particulier la manière dont une colonne se déduit des voisines devient obscure.

(iii) Cet example est si simple que le calcul et la factorisation du déterminant de (6) est immédiat. Nous donnerons plus bas (paragraphe 6) des indications sur la manière de procéder dans le cas général.

3. Préliminaires

Un *système d'équations algébriques*

$$\begin{cases} f_0(x_1, \ldots, x_n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ f_{k-1}(x_1, \ldots, x_n) = 0 \end{cases} \tag{8}$$

consiste en la donnée de k polynômes $f_0, \ldots, f_{k-1}$ en n variables et à coefficients dans un corps K. On dit alors que le système est *défini sur* K.

On peut associer au système (8) qui est défini sur K, l'anneau $B = K[x_1, \ldots, x_n]/(f_0, \ldots, f_{k-1})$ quotient de l'anneau des polynômes par l'idéal engendré par les f_i. Cet anneau est appelé *l'anneau affine du système.*

Si L est un corps, extension de K, on considèrera souvent l'anneau $B_L = L \otimes_K B = L[x_1, \ldots, x_n]/(f_0, \ldots f_{k-1})$. On appelle *solution dans L du système* (8) tout élément $(\xi_1, \ldots, \xi_n)$ de L^n tel que

$$f_i(\xi_1, \ldots, \xi_n) = 0 \quad \text{pour } i = 0, \ldots, k-1.$$

Théorème 3.1. *Les conditions suivantes sont équivalentes*:

(i) *Le système* (8) *n'a qu'un nombre fini de solutions dans une cloture algébrique* $\bar{K}$ *de* K.

(ii) *Pour toute extension* L *de* K, *le système* (8) *n'a qu'un nombre fini de solutions dans* L.

(iii) *L'anneau* B *est un* K*-espace vectoriel de dimension finie.*

Si ces conditions sont vérifiées, le nombre de solutions dans L est au plus égal au nombre de solutions dans $\bar{K}$ qui est lui-même au plus égal à la dimension de B comme K-espace vectoriel. En outre, si $(\xi_1, \ldots, \xi_n)$ est une solution, les ξ_i sont algébriques sur K. Enfin, les solutions dans $\bar{K}$ (et donc dans toute extension de $\bar{K}$) sont en bijection avec les idéaux maximaux de $B_{\bar{K}}$.

La démonstration de ce théorème en principe 'bien connu', comme celle du théorème 3.2 ci-dessous, est donnée en appendice. Le théorème 3.1 ne nous sera d'ailleurs guère utile, mais il nous semble indispensable pour la compréhension du théorème 3.2.

Appelons d_i le degré total du polynôme f_i (pour $i = 0, \ldots, k-1$). Nous pouvons associer à ce polynôme f_i le polynôme homogène

$$F_i(X_0, \ldots, X_n) = X_0^{d_i} f_i\left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}\right);$$

ainsi

$$f_i(x_1, \ldots, x_n) = F_i(1, x_1, \ldots, x_n).$$

Le système d'équations

$$\begin{cases} F_0(X_0, \ldots, X_n) = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ F_{k-1}(X_0, \ldots, X_n) = 0 \end{cases} \tag{9}$$

est le *système homogène* associé au système (8). Si $(\zeta_0, \ldots, \zeta_n)$ est une solution de (9), il en est de même de $(\lambda\zeta_0, \ldots, \lambda\zeta_n)$ pour tout λ. Il est donc naturel d'identifier deux telles solutions, et donc de considérer les solutions de (9) dans un corps L comme des éléments de l'espace projectif $\mathbb{P}_n(L)$; c'est ce que nous ferons lorsque nous considèrerons le nombre des solutions du système (9).

Si $(\xi_1, \ldots, \xi_n)$ est une solution de (8), alors $(1, \xi_1, \ldots, \xi_n)$ est une solution de (9); réciproquement, si $(\zeta_0, \ldots, \zeta_n)$ est une solution de (9) telle que $\zeta_0 \neq 0$, alors $(\zeta_1/\zeta_0, \ldots, \zeta_n/\zeta_0)$ est une solution de (8). Aussi, les solutions de (9) seront appelées *solutions projectives* du système (8); les solutions projectives pour lesquelles $\zeta_0 = 0$ sont dites *solutions à l'infini* du système (8); les autres solutions sont dites à *distance finie*.

On peut associer aux polynômes F_i l'anneau gradué $A = K[X_0, \ldots, X_n]/(F_0, \ldots, F_{k-1})$ quotient de l'anneau des polynômes en $n+1$ variables $X_0, \ldots, X_n$ par l'idéal engendré par les F_i. Si L est une extension de K, on pose

$$A_L = L \otimes_K A = L[X_0, \ldots, X_n]/(F_0, \ldots, F_{k-1}).$$

Il est immédiat que $B \simeq A/(X_0 - 1)A$ et que $B_L \simeq A_L/(X_0 - 1)A_L$. Cet anneau A est appelé *anneau du système* (9) ou *anneau gradué du système* (8).

Pour tout anneau ou module gradué X, (par exemple A, A_L ou $K[X_0, \ldots, X_n]$) on note X^d l'ensemble des éléments homogènes de degré d et on pose $X^+ = \oplus_{d>0} X^d$. On note enfin $\dim_K(V)$ la dimension du K-espace vectoriel V.

Théorème 3.2. *Les conditions suivantes sont équivalentes*:

(i) *Le système* (8) *n'a qu'un nombre fini de solutions projectives dans une clôture algébrique* $\bar{K}$ *de* K.

(ii) *Pour toute extension* L *de* K, *le système* (8) *n'a qu'un nombre fini de solutions projectives dans* L.

(iii) *Il existe un entier* D *tel que*

$$\dim_K(A^d) = \dim_K(A^D) \quad \textit{pour tout } d \geqslant D.$$

(iv) *Pour toute* (*resp. pour une*) *extension infinie* L *de* K, *il existe un entier* D' *et un élément* $y \in A_L^1$ *tels que la multiplication par* y *soit une surjection de* $A_L^{D'-1}$ *sur* $A_L^{D'}$.

Si ces conditions sont vérifiées, alors

(a) *La multiplication par* y *est une bijection de* A_L^d *sur* A_L^{d+1} *pour tout* $d \geqslant \max(D, D')$.

(b) *Le nombre de solutions projectives du système* (8) *est au plus* $\dim_K(A^D)$.

(c) *Si* L' *est la fermeture algébrique de* K *dans* L, *les solutions dans* $\mathbb{P}_n(L)$ *appartiennent en fait à* $\mathbb{P}_n(L')$.

(d) *Le système* (8) *vérifie les conditions du théorème* 3.1.

(e) *Le système* (9) *n'a aucune solution non triviale dans* $\bar{K}$ *si et seulement si* $A_D = 0$.

Il est manifeste que le théorème 3.2 est l'analogue projectif du théorème 3.1. L'intérêt du passage au système homogène (9) peut ne pas sembler évident; en fait cet intérêt provient de ce que l'on peut aisément calculer les entiers D et D'.

Théorème 3.3. *Si* d_i *désigne le degré commun des polynômes* F_i *et* f_i *et si ces polynômes sont rangés par ordre de degrés décroissants* $(d_0 \geqslant d_1 \geqslant \cdots \geqslant d_{k-1})$, *on peut prendre* $D = D' = d_0 + d_1 + \cdots + d_n - n$ *dans l'énoncé du théorème* 3.2. (*Si* $k \leqslant n$, *on pose* $d_i = 1$ *pour* $i \geqslant k$).

Dans les énoncés des théorèmes 3.1 et 3.2, les conditions équivalentes ne dépendent pas réellement du corps K; en effet, pour tout K-espace vectoriel V, on a $\dim_K(V) = \dim_L(L \otimes_K V)$. Nous avons fait intervenir un corps variable L (assertion (ii)) à cause de la situation suivante: Si K est le corps $\mathbb{Q}$ des rationnels, et si les conditions des théorèmes 3.1 ou 3.2 sont vérifiées, il revient au même de résoudre le système (8) dans $\mathbb{C}$ ou dans la clôture algébrique de $\mathbb{Q}$. Cette assertion est clairement fausse si les conditions équivalentes ne sont pas vérifiées (prendre une seule équation $x^2 - y^3 = 0$ qui admet comme solution $x = \pi^3$, $y = \pi^2$).

4. Réduction – Première partie

Supposons que le système (8) vérifie les conditions équivalentes du théorème 3.2, et considérons un élément quelconque

$$y = u_0 X_0 + u_1 X_1 + \cdots + u_n X_n,$$

de degré 1 de $K[X_0, \ldots, X_n]$. Si on ajoute l'équation

$$y = 0$$

au système (9), l'anneau A est remplacé par $A' = A/yA$. Deux cas peuvent se produire: s'il existe une solution du système (9) qui annule aussi y, on a $A'_D \neq 0$ (théorème 3.2(e)), ce qui signifie que la multiplication par y de A_{D-1} dans A_D n'est pas surjective. Si, au contraire, aucune solution de (9) n'annule y, on a $A'_D = 0$ et la multiplication par y est surjective. Ainsi l'étude de cette multiplication par y peut permettre de calculer les solutions des systèmes (8) et (9).

Plutôt que de faire varier les u_i dans K ou $\bar{K}$, nous allons introduire des indéterminées $U_0, \ldots, U_n$ et poser

$$L = U_0X_0 + \cdots + U_nX_n.$$

C'est un élément de $R[U_0, \ldots, U_n]$, avec $R = K[X_0, \ldots, X_n]$.

Pour simplifier les notations ultérieures, nous poserons toujours:

$$R = K[X_0, \ldots, X_n]$$

et

$$V_U = K[U_0, \ldots, U_n] \otimes_K V \quad \text{pour tout } K\text{-espace vectoriel } V;$$

nous identifierons les éléments v de V et leurs images $1 \otimes v$ dans V_U. Ainsi, par exemple,

$$L \in (R^1)_U \subset R_U,$$

puisque R est gradué (degré total en X). Nous considérerons toujours les U_i comme étant de degré zéro, c'est-à-dire que des notations telles que R^d_U, A^d_U signifieront $(R^d)_U$, $(A^d)_U$.

Ceci étant, considérons le diagramme commutatif

$$\begin{array}{ccc} R_U^{D-1} & \xrightarrow{\ L\ } & R_U^D \\ \downarrow{\scriptstyle p_U^{D-1}} & & \downarrow{\scriptstyle p_U^D} \\ A_U^{D-1} & \xrightarrow{\ L\ } & A_U^D \end{array}$$

où L désigne la multiplication par L et où les flèches verticales sont les projections canoniques. Ainsi, p_U^D est l'extension à R_U^D de la projection canonique $p^D : R^D \to A^D$, et le noyau de p^D est l'ensemble des $G_0F_0 + \cdots + G_{k-1}F_{k-1}$ où $G_i \in R^{D-d_i}$ pour tout i. Autrement dit (p^D, A^D) est le conoyau de l'application linéaire

$$\phi : R^{D-d_0} \times R^{D-d_1} \times \cdots \times R^{D-d_{k-1}} \to R^D$$

définie par $\phi(G_0, \ldots, G_{k-1}) = \sum G_iF_i$.

Il est facile d'écrire la matrice Φ de ϕ quand on prend pour base des K-espaces vectoriels qui apparaissent, la base formée par les monômes.

Exemple. Dans l'exemple du paragraphe 2, cette matrice Φ est la partie gauche du tableau (3).

Nous appellerons *réduction de Gauss* d'une matrice l'opération récursive suivante: si la matrice est nulle ou vide, ne rien faire, sinon choisir un élément non nul de cette matrice (le pivot), ajouter à toutes les lignes autres que celle du pivot un multiple de la ligne du pivot de manière à annuler l'élément correspondant de la colonne du pivot et faire une réduction de Gauss de la sous-matrice obtenue en supprimant la ligne et la colonne du pivot.

Ainsi une réduction de Gauss sur la matrice Φ permet de calculer une nouvelle base $e_1, \ldots, e_s$ de R^D telle que $e_1, \ldots, e_r$ soit une base de l'image de ϕ et que la matrice de ϕ sur cette nouvelle base ait la forme

$$\begin{array}{l} r \left\{ \right. \\ s-r \end{array} \left(\begin{array}{cccc|c} 1 & & & * & \\ & 1 & & & \\ & & 1 & & * \\ 0 & & & 1 & \\ \hline & & 0 & & \end{array} \right) = \Phi'$$

$\underbrace{\qquad\qquad}_{r}$

(à permutation près des colonnes et des lignes et à multiplication près des lignes par les inverses des pivots) où les $*$ désignent des coefficients quelconques. Il est clair que la restriction de p^D au sous-espace vectoriel engendré par $e_{r+1}, \ldots, e_s$ est un isomorphisme; en particulier $\dim_K(A_D) = s - r$; il en résulte que, si $r = s$, le système d'équations considéré n'a pas de solutions.

Appellons C la matrice de changement de base que cette réduction de Gauss nous a permis d'obtenir; on a $\Phi' = C\Phi$.

Remarque. Pour programmer cette réduction de Gauss et calculer la matrice C qui est seule utile dans la suite, il n'est pas utile d'écrire d'un seul coup la matrice Φ; il suffit d'écrire ses colonnes une après l'autre dans la même zone mémoire: initialisant C en matrice unité, on procède comme suit pour chaque colonne de Φ: écrire la colonne; la multiplier par l'ancienne valeur de C; chercher un pivot dans la colonne obtenue et dans les lignes où on n'a pas encore trouvé de pivot; s'il y a un pivot, les opérations sur les lignes qui annulent le reste de la colonne considérée correspondent à la multiplication à gauche par une matrice C_1; remplacer C par C_1C et passer à la colonne suivante. Cette remarque est utilisée dans l'algorithme explicite de [7].

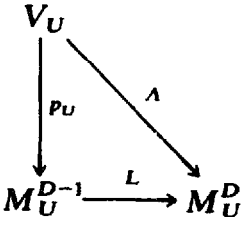
Revenant aux $K[U_0, \ldots, U_n]$-modules, il est facile d'écrire la matrice de $L: R_U^{D-1} \to R_U^D$ sur les bases constituées par les monômes, matrice que nous noterons

aussi L. Il est immédiat que CL est la matrice de L sur la nouvelle base de R_U^D calculée ci-dessus (toute base de R^D est aussi une base de R_U^D). Sur la base de A_U^D constituée par les images de $e_{r+1}, \ldots, e_s$, la matrice de $p_U^D L = L p_U^{D-1}$ est constituée par les $s-r$ dernières lignes de CL, et donc aisée à calculer.

Exemple. Dans l'exemple du paragraphe 2, la matrice L est la partie droite du tableau (3) et les $s-r$ dernières lignes de CL constituent la partie non triviale de la matrice (4).

A ce stade de la réduction nous devons réexprimer le problème afin de lui donner une forme plus générale qui nous permettra d'opérer récursivement.

Nous avons donc affaire à un A-module gradué M engendré par ses éléments de degré $<D$ (ici $M=A$); la composante M^{D-1} de M est exprimée comme quotient d'un K-espace vectoriel V (ici $V=R^{D-1}$), et l'on connait la matrice Λ du composé de la multiplication par L et de la projection $p: V \to M^{D-1}$;



Comme nous supposons vérifiées les conditions équivalentes du théorème 3.2, pour toute extension infinie K' de K, il existe y dans $A_{K'}^1$ tel que la multiplication par y soit une surjection de $M_{K'}^{D-1}$ sur $M_{K'}^D$ et une bijection de $M_{K'}^d$ sur $M_{K'}^{d+1}$ pour tout $d \geq D$. Cette condition sera appellée plus bas *condition* (Y)

Lemme 4.1. *Si $x \in M_{K'}^{D-1}$, si $z \in A_U^1$, et si la condition (Y) est vérifiée, on a:*

$$yx = 0 \Rightarrow zx = 0.$$

En raison de la commutativité de A, on a $yzx = zyx = 0$; comme $zx \in A_U \otimes M_{K'}^D$, la multiplication par y est injective, ce qui entraîne que $zx = 0$.

Proposition 4.1. *Si la condition (Y) est vérifiée, il existe une base de V définie par une matrice Γ à coefficients dans K, sur laquelle la matrice de Lp_U est de la forme*

$$\Lambda\Gamma = (\Lambda' \ 0)$$

où Λ' est une matrice carrée.

Exemple. Au paragraphe 2, $\Lambda\Gamma$ est la matrice (6).

Si K est infini, on peut prendre $K' = K$ et choisir la base de V de manière que le noyau de yp (qui est une application surjective) soit engendré par les derniers éléments de cette base. Le Lemme 4.1 implique alors que le noyau de Lp_U contient le

noyau de yp, ce qui démontre la proposition, compte tenu des dimensions de V, M^D et $\ker(yp)$.

Si K est fini, il faut utiliser le lemme 4.1 autrement: si $x \in M_{K'}^{D-1}$, alors $yx = 0$ implique $X_i x = 0$ pour tout i; autrement dit, $\ker(yp) \subset \bigcap \ker(X_i p)$. Il y a en fait égalité car y est combinaison linéaire (à coefficients dans K') des X_i. On a donc, puisque yp est surjectif,

$$\dim_{K'}(V_{K'}) = \dim_{K'}(\bigcap \ker(X_i p)) + \dim_{K'}(M_{K'}^D). \tag{10}$$

Comme tous ces espaces vectoriels sont, en fait, définis sur K, cette égalité est également vraie sur K, et il suffit de prendre une base de V dont les derniers éléments constituent une base de $\bigcap \ker(X_i p)$; comme Lp_U s'annule sur $\bigcap \ker(X_i p)$, l'égalité (10) entraine la proposition.

Théorème 4.1. *Soit r le nombre de lignes de Λ; supposons la condition (Y) vérifiée.*

(a) *Le rang de Λ est égal à r.*

(b) *L'idéal engendré par les déterminants $r \times r$ extraits de Λ est principal et engendré par un polynôme $G(U_0, \ldots, U_n)$ homogène et de degré r en $U_0, \ldots, U_n$.*

(c) *Si $M = A$ le polynôme G se décompose, sur une clôture algébrique $\bar{K}$ de K en produits de facteurs du premier degré. Le polynôme $\alpha_0 U_0 + \cdots + \alpha_n U_n$ est un tel facteur si et seulement si $(\alpha_0, \ldots, \alpha_n)$ est un zéro commun aux F_i.*

(a) résulte de la surjectivité de la multiplication par y : si $y = \sum y_i X_i$, la matrice de yp s'obtient en substituant les y_i aux U_i dans la matrice de Λ, et cette substitution ne peut que diminuer le rang.

(b) La multiplication par Γ (cf. proposition 4.1)) peut s'obtenir comme succession d'operations consistant soit à multiplier une colonne par un élément de K, soit à ajouter un multiple d'une colonne à une autre colonne. Ces opérations ne changent pas l'idéal engendré par les déterminants, ce qui montre que cet idéal est égal à l'idéal engendré par le déterminant G de Λ'.

(c) Pour démontrer la troisième assertion, adjoignons l'équation $u_o X_o + \cdots + u_n X_n = 0$ aux équations $F_i = 0$. Celà détermine un nouvel anneau $A' = A/(u_o X_o + \cdots + u_n X_n)$, et $A'^D = 0$ si et seulement si le nouveau système n'a pas de zéro commun. Mais $A'^D = A^D/(u_0 X_0 + \cdots + u_n X_n) A^{D-1}$ et $A'^D = 0$ signifie donc que la multiplication par $u_0 X_0 + \cdots + u_n X_n$ est surjective, c'est-à-dire que $G(u_o, \ldots, u_n) \neq 0$. Par ailleurs si les $(\alpha_{0,j}, \ldots, \alpha_{n,j})$ $(j = 1, \ldots, h)$ sont les zéros communs de l'ancien système, le nouveau système n'a pas de zéro commun si et seulement si

$$\prod_j (\alpha_{0,j} u_0 + \cdots + \alpha_{n,j} u_n) \neq 0.$$

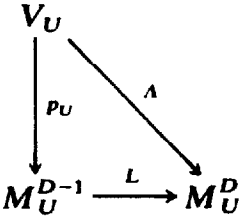
Autrement dit $\prod_j (\alpha_{0,j} U_0 + \cdots + \alpha_{n,j} U_n)$ et $G(U_0, \ldots, U_n)$ ont les mêmes zéros, ce qui donne le résultat à l'aide du théorème des zéros [5, théorème 33, par exemple].

Remarque. Nous verrons plus loin (paragraphe 7) que la multiplicité d'un facteur de G est égale à la multiplicité du zéro correspondant. Autrement dit, ce théorème permet non seulement de calculer les zéros, mais aussi de déterminer leur multiplicité. En particulier, on peut reconnaître les zéros multiples (singuliers) des autres.

5. Réduction – Deuxième partie

Au stade actuel de l'algorithme, nous ne connaissons pas l'élément y, ni a fortiori la matrice Γ, et il est très onéreux de calculer directement G comme PGCD d'une famille de déterminants. Aussi, allons-nous procéder à une suite de réductions récursives qui vont aboutir à obtenir la matrice Λ' sous forme de matrice triangulaire par blocs. L'intérêt de cette réduction, outre d'éviter de calculer des PGCD, est d'obtenir G sous une forme déjà partiellement factorisée. Cette méthode de réduction amènera à travailler sur des A-modules M qui seront obtenus comme quotients de sous-modules de A.

Revenons donc au diagramme



Les coefficients de la matrice Λ sont des polynômes homogènes du premier degré en U. Choisissons un indice i tel que U_i apparaisse dans Λ; par exemple supposons que U_0 apparaisse dans Λ et prenons $U_i = U_0$.

En procédant à une réduction de Gauss sur la matrice des coefficients de U_0, et en effectuant les mêmes opérations sur les lignes de Λ il est facile de calculer une matrice carrée Γ_1 à coefficients dans K telle que

$$\Gamma_1\Lambda = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}$$

où Λ_0 est une matrice triangulaire inversible, Λ_1, Λ_2, Λ_4 sont des matrices indépendantes de U_0 et où Λ_2 peut dépendre de U_0.

Exemple. Au paragraphe 2, la matrice $\Gamma_1\Lambda$ est la matrice (5). Cependant, pour avoir des matrices plus simples, nous avons effectué les réductions de Gauss du paragraphe 2 aussi bien sur les lignes que sur les colonnes ('diagonalisation'). C'est pourquoi la matrice Λ_0 est diagonale et Λ_2 ne dépend pas de U $(= U_0)$.

Deux cas sont à considérer:

Premier cas. Le nombre de lignes de Λ_3 et Λ_4 est nul, i.e.

$$\Gamma_1\Lambda = (U_0\Lambda_0 + \Lambda_1 \; \Lambda_2).$$

Décomposons Λ_2 en $\Lambda_2 = U_0\Lambda_2' + \Lambda_2''$ où Λ_2' est à coefficients dans K et Λ_2'' à coefficients dans $K[U_1, \ldots, U_n]$. Posons

$$\Gamma_2 = \begin{pmatrix} I_1 & -\Lambda_0^{-1}\Lambda_2' \\ 0 & I_2 \end{pmatrix}$$

où I_1 et I_2 sont des matrices unités de dimensions convenables. On a

$$\Gamma_1\Lambda\Gamma_2 = (U_0\Lambda_0 + \Lambda_1 \quad \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2').$$

Lemme 5.1. $\Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda'^2 = 0$.

Utilisant les notations de la proposition 4.1, on a

$$(U_0\Lambda_0 + \Lambda_1 \; \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2') = \Gamma_1(\Lambda' \; 0)\Gamma^{-1}\Gamma_2 = (\Gamma_1\Lambda' \; 0)\Gamma^{-1}\Gamma_2$$

et il existe donc deux matrices Δ_1 et Δ_2 à coefficients dans K telles que

$$U_0\Lambda_0 + \Lambda_1 = \Gamma_1\Lambda'\Delta_1 \quad \text{et} \quad \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2' = \Gamma_1\Lambda'\Delta_2.$$

Substituant 1 à U_0 et 0 aux autres U_i, il vient

$$\Lambda_0 = \Gamma_1\Lambda'(1, 0, \ldots, 0)\Delta_1 \quad \text{et} \quad 0 = \Gamma_1\Lambda'(1, 0, \ldots, 0)\Delta_2$$

Ainsi $\Gamma_1\Lambda'(1, 0, \ldots, 0)$ est inversible et $\Delta_2 = 0$, ce qui démontre le lemme.

Corollaire. *G est le déterminant des colonnes non nulles de $\Gamma_1\Lambda\Gamma_2$.*

Deuxième cas. Le nombre de lignes de Λ_3 et Λ_4 est $r \neq 0$.

Il y a deux manières de traiter cette situation: on peut montrer que la matrice $(\Lambda_3 \; \Lambda_4)$ est la matrice d'une application $\Lambda : V_U \to M_U^D$ correspondant à un module M vérifiant la condition (Y), puis calculer la matrice Γ telle que $(\Lambda_3 \; \Lambda_4)\Gamma = (\Lambda' \; 0)$ pour une certaine matrice carrée Λ'; c'est possible, récursivement, car le nombre de lignes de $(\Lambda_3 \; \Lambda_4)$ est strictement inférieur à celui de Λ. On a alors

$$\Gamma_1\Lambda\Gamma = \begin{pmatrix} * & \Lambda'' \\ \Lambda' & 0 \end{pmatrix}.$$

On peut montrer que Λ'' est la matrice d'une autre application $\Lambda : V_U \to M_U^D$ correspondant à un autre module M satisfaisant la condition (Y). Une autre récursion permet donc de calculer une matrice Γ' tel que $\Lambda''\Gamma' = (\Lambda''' \; 0)$ où Λ''' est également une matrice carrée. Il est clair que le polynôme cherché G est le produit des déterminants de Λ' et de Λ'''.

L'inconvénient de cette méthode est que la récursion suit un arbre binaire et sa programmation est donc plus compliquée que la méthode que nous allons exposer. Celle-ci consiste également à interpréter $(\Lambda_3 \; \Lambda_4)$ comme la matrice d'une application $\Lambda : V_U \to M_U^D$, mais, au lieu de mettre cette matrice sous la forme $(\Lambda' \; 0)$, on lui applique simplement la discussion du début: on choisit une indéterminée U_i,

par exemple U_1, et des matrices $\Gamma_1^1, \Lambda_0^1, \ldots, \Lambda_4^1$ telles que

$$\Gamma_1^1(\Lambda_3 \ \Lambda_4) = \begin{pmatrix} U_1\Lambda_0^1 + \Lambda_1^1 & \Lambda_2^1 \\ \Lambda_3^1 & \Lambda_4^1 \end{pmatrix}$$

que Λ_0^1 soit triangulaire inversible et que $\Lambda_1^1, \Lambda_3^1, \Lambda_4^1$ soient indépendants de U_1 (et aussi, évidemment, de U_0). Si le nombre de lignes de $(\Lambda_3^1 \ \Lambda_4^1)$ n'est pas nul, on peut recommencer, écrire

$$\Gamma_1^2(\Lambda_3^1 \ \Lambda_4^1) = \begin{pmatrix} U_2\Lambda_0^2 + \Lambda_1^2 & \Lambda_2^2 \\ \Lambda_3^2 & \Lambda_4^2 \end{pmatrix}$$

et itérer jusqu'à obtenir un entier k tel que

$$\Gamma_1^k(\Lambda_3^{k-1} \ \Lambda_4^{k-1}) = (U_k \ \Lambda_0^k + \Lambda_1^k \ \Lambda_2^k).$$

Si on pose $\Lambda_2^k = U_k\Lambda_2' + \Lambda_2''$ avec Λ_2'' indépendant de U_k, et

$$\Gamma_2 = \begin{pmatrix} I_1 & -(\Lambda_0^k)^{-1}\Lambda_2' \\ 0 & I_2 \end{pmatrix},$$

le lemme 5.1 montre que

$$\Gamma_1^k(\Lambda_3^{k-1} \ \Lambda_4^{k-1})\Gamma_2 = (U_k\Lambda_0^k + \Lambda_1^k \ \ 0).$$

Posons, pour tout i,

$$\Gamma^i = \begin{pmatrix} I & 0 \\ 0 & \Gamma_1^i \end{pmatrix}$$

où I est une matrice identité de dimension convenable. On a donc

$$\Gamma^k\Gamma^{k-1}\cdots\Gamma^1\Gamma_1\Lambda\Gamma_2 = \begin{pmatrix} \Lambda^2, & \Lambda^1 \\ U_k\Lambda_0^k + \Lambda_1^k & 0 \end{pmatrix}.$$

Aussi l'idéal engendré par les déterminants de Λ est le produit du déterminant de $U_k\Lambda_0^k + \Lambda_1^k$ par l'idéal engendré par les déterminants de Λ^1. En appliquant la discussion précédante à Λ^1 et en itérant, on obtient le théorème suivant, sous réserve que l'on puisse appliquer le Lemme 5.1 chaque fois que nécessaire.

Exemple. Au paragraphe 2, les matrices Λ_3^1 et Λ_4^1 n'ont aucune lignes (i.e. $k = 1$), la matrice $U_k\Lambda_0^k + \Lambda_1^k$ est simplement $-V + W$ et la matrice Λ^1 est simplement la matrice $I_0\Lambda_0 + \Lambda_1$ de l'étape précédente, bordée de deux colonnes nulles. Sa réduction est donc déjà faite.

Théorème 5.1. *L'algorithme précédemment décrit permet de calculer le polynôme G comme produit des déterminants d'un certain nombre de matrices carrées de la forme $U_k\Lambda_0^k + \Lambda_1^k$ où Λ_0^k est une matrice triangulaire inversible à coefficients dans K et Λ_1^k une matrice carrée à coefficients linéaires et homogènes en $U_{k+1}, U_{k+2}, \ldots, U_n$.*

Remarque. La procédure REDUC de l'appendice A.4 est exactement une traduction de l'algorithme précédent, à un détail près: dans REDUC, on a systématiquement multiplié à droite les matrices qui interviennent par les matrices

$$\begin{pmatrix} (\Lambda_0^i)^{-1} & -(\Lambda_0^i)^{-1}\Lambda_2' \\ 0 & I \end{pmatrix},$$

ce qui allonge le temps de calcul, mais simplifie légèrement la programmation.

Pour terminer la démonstration du théorème 5.1, il faut donc interpréter la signification des matrices $(\Lambda_3\ \Lambda_4)$ et Λ^1. Pour cela, désignons par $\varepsilon_1, \ldots, \varepsilon_s$ et $e_1, \ldots, e_t$ les bases de V et M^D sur lesquelles la matrice de Lp_U est

$$\Gamma_1\Lambda = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}.$$

En substituant 1 à U_0 et 0 aux autres U_i, on voit que l'image de la multiplication par X_0 est engendrée par $e_1, \ldots, e_r$ où r est la dimension de Λ_0. Ceci montre que $(\Lambda_3\ \Lambda_4)$ est la matrice de l'application composée

$$p_1 : V_U \xrightarrow{Lp_U} M_U^D \to (M/X_0M)_U^D.$$

Posons $M' = M/X_0M$; il est facile de vérifier que M' vérifie la condition (Y) et que p_1 est l'application composée $L'p'_U$ où L' est la multiplication par L dans M' et p' l'application composée de p et de la surjection canonique $M^{D-1} \to M^{D-1}/X_0M^{D-2}$. Ceci montre que le lemme 5.1 s'applique bien à la matrice $(\Lambda_3\ \Lambda_4)$.

En itérant ce qui précède, on voit que $(\Lambda_3^{k-1}\ \Lambda_4^{k-1})$ correspond à l'application

$$V_U \to (M/X_0M + \cdots + X_{k-1}M)_U^D$$

et que $M^D = X_0M^{D-1} + \cdots + X_kM^{D-1}$. Désignons par $e'_1, \ldots, e'_s$ et $\varepsilon'_1, \ldots, \varepsilon'_t$ les bases de V et M^D sur lesquelles la matrice de Lp_U est

$$\begin{pmatrix} \Lambda^2 & \Lambda^1 \\ U_k\Lambda_0^k + \Lambda_1^k & 0 \end{pmatrix}$$

et appelons r la dimension de $U_k\Lambda_0^k + \Lambda_1^k$. Les images de $\varepsilon'_{t-r+1}, \ldots, \varepsilon'_t$ constituent une base de $(M/X_0M + \cdots + X_{k-1}M)^D$ et les images de $e'_{s-r+1}, \ldots, e'_s$ dans ce module sont nulles.

Il en résulte que $X_0M^{D-1} + \cdots + X_{k-1}M^{D-1}$ est de dimension $t-r$ et a pour base $\varepsilon'_1, \ldots, \varepsilon'_{t-r}$. Appelons V' le sous-espace de V engendré par $e'_{r+1}, \ldots, e'_s$; l'image de V'_U par Lp_U est contenue dans $X_0M_U^{D-1} + \cdots + X_{k-1}M_U^{D-1}$. Posons enfin $M'' = Ap(V') + X_0M + \cdots + X_{k-1}M$. Comme les multiplications par X_i et par y commutent, on voit facilement que M'' vérifie la condition (Y), que la restriction p' de p à V' a son image dans M''^{D-1} et que la matrice de Lp' est Λ^1. L'application $p' : V' \to M''^{D-1}$ est surjective: raisonnons d'abord sur une extension infinie K' de K;

comme $Lp_U(V'_U) \subset M''^D_U$, on obtient, en substituant les coefficients de y aux U_i, l'inclusion $yp_{K'}(V'_{K'}) \subset M''^D_{K'}$; il en résulte, par passage au quotient, une application surjective de $V_{K'}/V'_{K'}$ sur $M^D_{K'}/M''^D_{K'}$, qui est bijective car il y a égalité des dimensions; on en déduit que l'application $p_{K'}$ induit une injection de $V_{K'}/V'_{K'}$ dans $M^{D-1}_{K'}/M''^{D-1}_{K'}$, ce qui entraîne que $p'_{K'} : V'_{K'} \to M''^{D-1}_{K'}$ est une surjection et qu'il en est de même de $p' : V' \to M''^{D-1}$.

Ce qui précède montre que la méthode décrite pour réduire la matrice A s'applique à la matrice A^1 et termine la démonstration du théorème 5.1.

Remarque. La validité de l'autre méthode de réduction esquissée plus haut se démontre de même, les matrices A' et A'' correspondant alors aux modules M/X_0M et $X_0M + p'(V')$. En fait, cette autre méthode s'est averée meilleure [13].

6. Calculs finaux

6.1. Cas où les conditions du théorème 3.2 ne sont pas vérifiées

Si le système possède une infinité de solutions projectives, le rang de la matrice A que l'on a réduite au paragraphe précédent est inférieur à son nombre de lignes et les modules qui apparaissent ne vérifient pas tous la condition (Y). Au cours du déroulement de l'algorithme, cette situation se révèle donc nécessairement par au moins un des deux faits suivants:

(a) Apparition d'une matrice $(A^i_3\ A^i_4)$ qui est nulle.

(b) Non validité du lemme 5.1 lors d'une de ses utilisations.

Ces deux faits se testent en comparant à zéro certaines matrices qui apparaissent pendant l'exécution de l'algorithme. Il en résulte que l'algorithme permet de vérifier si les conditions équivalentes du théorème 3.2 sont satisfaites.

Remarque. Le fait (a) est automatiquement testé au cours de l'exécution; il n'en est pas de même du fait (b); si l'on est certain de la validité des conditions équivalentes du théorème 3.2, on gagne donc du temps d'exécution en ne testant pas (b).

Il est possible que, si le système a une infinité de solutions, le fait (a) ait toujours lieu. Si tel était le cas, le gain de temps précédent serait toujours possible.

6.2. Solutions ayant une coordonnée nulle

L'algorithme décrit au paragraphe précédent donne les solutions $(a_0, \ldots, a_n)$ comme facteurs linéaires $a_0U_0 + \cdots + a_nU_n$ de déterminants de matrices de la forme $U_kA^k_0 + A^k_1$ où A^k_0 est une matrice triangulaire inversible à coefficients dans K et A^k_1 une matrice carrée à coefficients linéaires et homogènes relativement à $U_{k+1}, \ldots, U_n$. Les facteurs linéaires d'un tel déterminant sont tous de la forme

$$a_kU_k + \cdots + a_nU_n$$

avec $a_k \neq 0$. Les solutions correspondantes vérifient donc toutes $a_0 = a_1 = \cdots = a_{k-1} = 0$ et $a_k \neq 0$. Parmi ces solutions certaines peuvent vérifier $a_i = 0$ pour un certain $i > k$. Il est facite de les séparer des autres: il suffit de modifier l'ordre

$$U_k, U_{k+1}, \ldots, U_n$$

dans lequel on considère les U_i en

$$U_i, U_k, U_{k+1}, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n,$$

puis d'appliquer l'algorithme décrit au paragraphe 5 à la matrice $U_k \Lambda_0^k + \Lambda_1^k$.

Cette séparation des solutions a l'intérêt de réduire la taille des déterminants à calculer et peut donc simplifier fortement la fin de la résolution.

L'algorithme sépare en particulier les solutions à l'infini ($a_0 = 0$) des autres. Cela est particulièrement intéressant, car l'homogénéisation introduit souvent de nombreuses solutions parasites à l'infini.

6.3. Solutions sur un hyperplan particulier

Il peut arriver que, pour des raisons géométriques, il soit vraisemblable que certaines solutions appartiennent à un hyperplan particulier d'équation

$$b_0 X_0 + \cdots + b_n X_n = 0;$$

cela signifie que $a_0 b_0 + \cdots + a_n b_n = 0$.

Supposons, par exemple, que b_0 ne soit pas nul et effectuons le changement linéaire de variables

$$U_i' = U_i + (b_i/b_0) U_0 \quad \text{pour } i > 0, \qquad U_0' = U_0.$$

Le polynôme $a_0 U_0 + \cdots + a_n U_n$ devient alors

$$\frac{1}{b_0} \sum_{i=0}^{n} a_i b_i U_0' + a_1 U_1' + \cdots + a_n U_n'.$$

Ceci donne un moyen de séparer les solutions vérifiant $\sum a_i b_i = 0$ des autres: on effectue le changement de variables ci-dessus dans la matrice

$$U_k \Lambda_0^k + \Lambda_1^k$$

et on applique l'algorithme du paragraphe 5 à la matrice obtenue. Ceci conduit à des matrices du type $U_i' \Lambda_0^i + \Lambda_1^i$; celles de ces matrices pour lesquelles $i \neq 0$ correspondent aux solutions telles que $\sum a_i b_i = 0$.

6.4. Calcul des solutions

Il y a un cas où le calcul des solutions correspondant à

$$U_k \Lambda_0^k + \Lambda_1^k$$

est immédiat: c'est quand cette matrice est de dimension 1. Son unique coefficient

$$a_k U_n + \cdots + a_n U_n$$

est alors égal à son déterminant et donne donc sans calcul la solution

$$(a_k, \ldots, a_n).$$

Dans le cas général, la méthode consistant à calculer le déterminant de $H = U_k \Lambda_0^k + \Lambda_1^k$ et à factoriser le polynôme obtenu est onéreuse. Il est plus rapide de procéder comme suit: on peut supposer $\Lambda_1^k \neq 0$; sinon le déterminant cherché est simplement une puissance de U_k. Supposons donc, par exemple que U_{k+1} apparaisse dans Λ_1^k, et substituons 1 à U_{k+1} et 0 à $U_{k+2}, \ldots, U_n$; on obtient une matrice

$$\bar{H} = U_k \Lambda_0^k + \bar{\Lambda}_1^k$$

où Λ_0^k et Λ_1^k sont des matrices scalaires. Si $G(U_k, \ldots, U_n)$ est le déterminant de H et $\bar{G}(U_k) = G(U_k, 1, 0, \ldots, 0)$ celui de $\bar{H}$, il est immédiat que, pour tout facteur

$$a_k U_k + \cdots + a_n U_n$$

de G, on a $a_k \neq 0$ et $-a_{k+1}/a_k$ est une racine de $\bar{G}$. Comme les solutions et les facteurs de G sont définis à multiplication par un élément de K près, on peut supposer $a_k = 1$; il en résulte que les coefficients a_{k+1} des solutions sont les racines de l'équation $\bar{G}(-U_k) = 0$.

Supposons pour l'instant cette équation résolue, et soit a_{k+1} une de ses racines. Les solutions correspondantes sont sur l'hyperplan d'équation

$$a_{k+1} X_k - X_{k+1}.$$

L'argument décrit au paragraphe 6.3 ci-dessus permet, à l'aide d'une nouvelle application de l'algorithme du paragraphe 5, de séparer les solutions $(1, a_{k+1}, \ldots)$ des autres. Si, en outre il n'y a qu'une solution de la forme $(1, a_{k+1}, \ldots)$, ce qui est le cas le plus fréquent, ses coefficients sont obtenus comme fonctions rationnelles de a_{k+1}. S'il y a plusieurs solutions de la forme $(1, a_{k+1}, \ldots)$, l'algorithme du paragraphe 5 peut conduire à des matrices $U_k \Lambda_0^k + \Lambda_1^k$ de dimension supérieure à 1; mais leur dimension est inférieure à celle de la matrice dont on est parti; en faisant jouer successivement à $U_{k+2}, \ldots, U_n$ le rôle joué par U_{k+1} dans l'argument précédent, on finit par calculer tous les coefficients des solutions $(1, a_{k+1}, \ldots)$.

Il reste à voir comment calculer les racines de $\bar{G}(-U_k)$. Une première méthode consiste à calculer ce polynôme, par exemple par interpolation, et à calculer ses racines par n'importe quelle méthode classique. Une seconde méthode consiste à remarquer que les racines de $\bar{G}(-U_k)$ sont exactement les valeurs propres de $\bar{\Lambda}_1^k (\Lambda_0^k)^{-1}$ (rappelons que Λ_0^k est triangulaire inversible). Les méthodes de calcul des valeurs propres permettent donc de résoudre directement l'équation $\bar{G}(-U_k) = 0$ sans calculer le polynôme $\bar{G}$.

Si l'on s'intéresse à la structure algébrique de l'ensemble des solutions, il y a toutefois intérêt à calculer $\bar{G}$ et à le factoriser en facteurs irréductibles et unitaires. Si

$f(U_k)$ est un tel facteur, on peut travailler dans le corps $K(a_{k+1}) = K[X]/(f(-X))$; la poursuite de la résolution donne alors les solutions, le plus souvent rationnellement, en fonction de a_k; en outre, les solutions conjuguées sont regroupées.

7. Multiplicité

Il est fréquent que certaines des solutions du système considéré soient des solutions multiples; par exemple l'intersection d'un cercle et d'une tangente est un point double. Ces solutions multiples ont souvent une origine géométrique différente des solutions simples et doivent être considérées comme irrégulières. Il est donc utile de les reconnaître et d'en calculer éventuellement la multiplicité. L'algorithme que nous avons décrit le permet sans calculs supplémentaires, comme l'affirme le théorème suivant dont la démonstration est l'objet de ce paragraphe.

Théorème 7.1. *Dans l'énoncé du théorème* 4.1, *la multiplicité d'un facteur* $a_0U_0 + \cdots + a_nU_n$ *de* G *est égale à la multiplicité de la solution* $(a_0, \ldots, a_n)$ *correspondante.*

Pour que cet énoncé ait un sens, il faut une définition précise de la multiplicité d'une solution $(a_0, \ldots, a_n)$. Par définition, cette multiplicité sera la multiplicité de l'anneau A en l'idéal premier $\mathfrak{p}$ engendré par les $a_iX_j - a_jX_i$ (voir proposition A.3 et [4, p. 51]), c'est-à-dire la longueur de l'anneau artinien $A_{\mathfrak{p}}$.

Effectuons un changement linéaire et homogène de variables, de manière que, relativement aux nouvelles variables, la solution $(a_0, \ldots, a_n)$ devienne $(0, \ldots, 0, 1)$: il suffit de prendre $X_i' = X_i - (a_i/a_n)X_n$ pour $i < n$, $X_n' = X_n$, $U_n' = a_0U_0 + \cdots + a_nU_n$ et $U_i' = U_i$ pour $i \neq n$ (on a supposé $a_n \neq 0$, ce que l'on peut toujours faire après permutation éventuelle des variables).

Lorsqu'on applique l'algorithme du paragraphe 5 le polynôme G correspondant à un module M est décomposé en un produit d'un polynôme G' correspondant à un module $M' = M/(X_0'M + \cdots + X_{k-1}'M)$ et d'un polynôme G'' correspondant à un module $M'' = X_0'M + \cdots + X_{k-1}'M + p(V')$ appelons M''' le module $X_0'M + \cdots + X_{k-1}'M$. De la suite exacte

$$0 \to M''' \to M \to M' \to 0,$$

on déduit la relation

$$\text{mult}(\mathfrak{p}, M) = \text{mult}(\mathfrak{p}, M''') + \text{mult}(\mathfrak{p}, M')$$

entre les multiplicités ($\text{mult}(\mathfrak{p}, M)$ est la longueur de $M_{\mathfrak{p}}$). Mais, comme $\mathfrak{p}$ est engendré par $X_0', \ldots, X_{n-1}'$, l'élément X_n' est inversible dans $A_{\mathfrak{p}}$; il en résulte que si $z \in p(V')$, l'image $z/1$ de z dans $M_{\mathfrak{p}}''$ est égale à $X_n'z/X_n'$, ce qui montre que $M_{\mathfrak{p}}'' = M_{\mathfrak{p}}'''$ (car $Lp(V') \subset M_U'''$) et que M'' et M''' ont même multiplicité.

Il résulte donc de ce qui précède qu'il suffit de démontrer le théorème lorsque G est le déterminant d'une matrice de la forme $U_kA_0^k + A_1^k$ correspondant à un certain module M, et lorsque la solution est $(0, \ldots, 0, 1)$.

Si $k \neq n$ G n'a aucun facteur égal à X_n; par ailleurs, X_k est nilpotent et X_n inversible dans l'anneau artinien $A_{\mathfrak{p}}$; comme la multiplication par X_k de M^{d-1} dans M^d est surjective pour $d \geq D$ on a, pour tout z de M et pour tout l,

$$X_n^{D+l} z = X_k^l t;$$

si l est assez grand, $X_k^l = 0$ dans $A_{\mathfrak{p}}$, ce qui entraîne que $M_{\mathfrak{p}} = 0$, c'est-à-dire $\text{mult}(\mathfrak{p}, M) = 0$.

Si $k = n$, on a $\Lambda_1^n = 0$, et $G = \lambda U_n^r$, avec $\lambda \in K$ et $r = \text{rang}(\Lambda_0^n)$. Il existe, d'autre part une suite $0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_i = M$ de sous-modules gradués de M tels que, pour tout i, $M_{i+1}/M_i \simeq (A/\mathfrak{p}_i)(l_i)$ où $\mathfrak{p}_i$ est un idéal premier gradué de A et $(A/\mathfrak{p}_i)(l_i)$ est défini par $(A/\mathfrak{p})(l_i)^d = (A/\mathfrak{p})^{d+l_i}$ [4, paragraphe I.7.4]. Comme $\dim_K (M^d)$ est indépendant de d pour d assez grand, il en est de même $\dim_K (M_{i+1}^d/M_i^d)$. Si $\mathfrak{p}_i$ est l'idéal $\mathfrak{m}$ engendré par tous les X_i, on a $A/\mathfrak{p}_i \simeq K$ et cette dimension est nulle; si $\mathfrak{p}_i$ est un idéal premier différent de $\mathfrak{m}$ et $\mathfrak{p}$, il existe $z \in A_1$ tel que $z \in \mathfrak{p}$ et $z \notin \mathfrak{p}_i$; posons $z = \sum_{i=0}^{n-1} z_i X_i$; en substituant les z_i aux U_i dans $U_n \Lambda_0^n$, on voit que la multiplication par z annule M^{D-1}, et donc aussi $(M_{i+1}/M_i)^d$ pour $d \geq D-1$; ceci implique que $z^D M_{i+1}/M_i = 0$, ce qui est contraire aux hypothèses, car z ne divise pas zéro dans M_{i+1}/M_i. Ainsi, si $\mathfrak{p}_i \neq \mathfrak{m}$, on a $\mathfrak{p}_i = \mathfrak{p}$, et, comme $A/\mathfrak{p} \simeq K[X_n]$, on a $\dim_K (M_{i+1}^d/M_i^d) = 1$ pour d assez grand.

On a donc montré que $\dim_K M^D$ est exactement le nombre d'indices i tels que $\mathfrak{p}_i = \mathfrak{p}$. Il est facile de vérifier que l'anneau artinien $A_{\mathfrak{p}}$ contient le corps $K(X_n)$ et que $A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}} \simeq K(X_n)$; ceci montre que

$$\text{mult}(\mathfrak{p}, M) = \dim_{K(X_n)} (M_{\mathfrak{p}}) = \sum_i \dim_{K(X_n)} (M_{i+1}/M_i)_{\mathfrak{p}}.$$

Mais $(M_{i+1}/M_i)_{\mathfrak{p}} \simeq (A/\mathfrak{p}_i) \otimes A_{\mathfrak{p}}$; ce module est donc nul si $\mathfrak{p}_i = \mathfrak{m}$ et isomorphe à $K(X_n)$ si $\mathfrak{p}_i = \mathfrak{p}$. Ceci montre donc que $\text{mult}(\mathfrak{p}, M) = \dim_K M^D = r$ et termine la démonstration du théorème 7.1.

8. Complexité

La complexité de l'algorithme que nous avons décrit dépend évidemment du corps K sur lequel on travaille. Il y a trois cas importants que conduisent à des estimations de complexité différentes:

(a) La complexité des opérations de $\bar{K}$ est constante; c'est le cas lorsque l'on travaille dans le corps des complexes.

(b) La complexité des opérations de K est constante mais pas celle des opérations de $\bar{K}$; c'est le cas lorsque K est un corps fini.

(c) La complexité des opérations de K n'est pas constante; c'est le cas lorsque K est le corps des rationnels.

Par ailleurs, la résolution complète comprend la résolution d'une (ou plusieurs) équations en une inconnue. Cette complexité dépend également du corps sur lequel

on travaille. Si, en particulier, le corps K est $\mathbb{Q}$ ou un corps fini, cette résolution est, en fait, une factorisation, et la complexité totale de l'algorithme décrit dépend de celle d'une telle résolution-factorisation.

Enfin, la complexité de notre algorithme dépend beaucoup de la structure algébrique et géométrique de l'ensemble des solutions.

D'un autre côté, la difficulté du problème dépend de nombreux paramètres, à savoir le nombre d'équations k, le nombre de variables n et la suite des degrés $d_0, \ldots, d_{k-1}$.

Tout ceci fait qu'une étude complète de la complexité sortirait du cadre de cet article, et que nous nous limiterons au cas particulier où toutes les équations ont le même degré d,

Lemme 8.1. *Si toutes les équations ont le même degré d, le nombre L des lignes et le nombre C des colonnes de la matrice Φ du paragraphe 4 vérifient*

$$L \leqslant (ed)^n \quad et \quad C \leqslant k(ed)^n.$$

On a en effet $D = (n+1)d - n$,

$$L = \frac{(D+1)(D+2)\cdots(D+n)}{n!} \leqslant \frac{((n+1)d)^n}{n!} \leqslant (ed)^n$$

et

$$C = k\frac{(D+1-d)(D+2-d)\cdots(D+n-d)}{n!} \leqslant kL.$$

Proposition 8.1. *L'étape de l'algorithme décrite au début du paragraphe 4 nécessite $O(k(ed)^{3n})$ opérations de K et la mise en mémoire simultanée d'au plus $O((ed)^{2n})$ éléments de K.*

Cela résulte immédiatement de la réduction de Gauss qui est effectuée, compte tenu que l'on n'a pas besoin de stocker la matrice Φ, mais seulement la matrice C.

Proposition 8.2. *S'il y a N solution comptées avec leur multiplicité, l'étape de l'algorithme décrite au paragraphe 5 nécessite la mise en mémoire simultanée d'au plus $O(ne^n d^n N)$ éléments de K; le nombre d'opérations de K est $O(ne^n d^n N^3)$, mais seulement $O(ne^n d^n N^2)$ s'il n'y a pas de solution à l'infini.*

La matrice A du paragraphe 4 a donc N lignes; son nombre de colonnes est majoré par $(ed)^n$ et chacun de ses coefficients est constitué de $n+1$ éléments de K. Pour écrire cette matrice on a besoin également des N dernières lignes de la matrice C qui a moins de $(ed)^n$ colonnes. Il en résulte que l'écriture de A nécessite au plus $(n+2)e^n d^n N$ mémoires; il est facile de vérifier que le calcul de $A = CL$ ne nécessite aucune opération de K.

La réduction de A consiste en une suite de réductions de Gauss; lorsque le nombre de pivots trouvés est égal au nombre de lignes de A, on sait extraire de A une sous-matrice carrée. En général, il n'y a pas de solution à l'infini, et l'étape est terminée après un nombre de pivotages égal au nombre de lignes N de A. Ceci conduit donc à, au plus $O(N^2(n+1)(ed)^n)$ opérations. S'il y a beaucoup de solutions sur les hyperplans coordonnées, il peut y avoir jusqu'à N matrices carrées $U_kA_0^k+A_1^k$ isolées par l'algorithme du paragraphe 5, ce qui peut conduire jusqu'a $\frac{1}{2}N(N+1)$ pivots, et donc $O(ne^nd^nN^3)$ opérations.

Remarque. Cette estimation est très grossière, car la séparation d'une sous-matrice carrée ne détruit pas complètement le pivotage qui a déjà été effectué sur le reste de la matrice.

Théorème 8.1. *Considérons un système* (8) *dont toutes les équations sont de degré au plus d et qui possède N solutions simples dont toutes les coordonnées sont distinctes et non nulles (situation générale). La résolution de ce système selon l'algorithme décrit nécessite au plus*

$+O(ke^{3n}d^{3n})$ *opérations de K*

$+O(nN^4)$ *opérations dans des extensions de degré N de K ou dans la clôture algébrique de K*

+ Le calcul des valeurs propres d'une matrice $N\times N$ ou la résolution d'une équation de degré N en une inconnue et $O(N^{3.8})$ opérations de K.

Comme $N\leqslant d^n$ (théorème de Bezout), on a $nN^2e^nd^n<ke^{3n}d^{3n}$ et les deux premières étapes, ensemble, nécessitent $O(ke^{3n}d^{3n})$ opérations. Le calcul des premières coordonnées des solutions nécessite, soit le calcul des valeurs propres d'une matrice $N\times N$, soit le calcul de son polynôme caractéristique et sa factorisation. Le calcul du polynôme par interpolation nécessite celui de N déterminants de K, soit moins de $O(N^{3.8})$ opérations, la complexité de l'interpolation étant négligeable.

Connaissant la première coordonnée d'une solution, on peut appliquer de nouveau l'algorithme du paragraphe 5 pour obtenir toutes les autres en $O(nN^3)$ opérations; les N solutions nécessitent donc, ensemble, $O(nN^4)$ opérations dans des extensions de K, ce qui démontre le théorème.

Corollaire. *Si les solutions vérifient les hypothèses du théorème* 8.1, *et si $k\leqslant 2n$ et $d\geqslant 3$, l'algorithme est polynomial relativement au nombre maximal théorique de solutions.*

Ce nombre maximal est d^n (théorème de Bezout). Il suffit donc de vérifier que les opérations dans une extension algébrique de degré d^n de K sont polynomiales en terme d'opérations de K, et que la resolution d'une équation algébrique l'est également. Le premier point est facile. Le second dépend beaucoup du corps K considéré: si K est le corps des réels ou des complexes, il en est bien ainsi, à condition

de se limiter à une précision donnée. Il en est de même si K est un corps fini, mais le résultat correspondant pour le corps des rationnels est encore conjectural [3].

L'estimation qui précède de la complexité de l'algorithme considéré, bien que très partielle permet de le comparer aux algorithmes basés sur le résultant:

Proposition 8.3. *Sous les hypothèses du théorème* 8.1, *les algorithmes basés sur le résultant nécessitent au moins* $d^{2^{n-1}}$ *opérations de* K *et la mémorisation d'au moins* $d^{2^{n-1}}$ *éléments de* K.

L'utilisation du résultant pour éliminer une variable double les degrés des polynômes. Cette méthode conduit donc à calculer et manipuler un polynôme de degré $d^{2^{n-1}}$.

Remarques. (1) Ce résultat montre que, pour les grandes valeurs de n, notre algorithme est plus rapide que ceux basés sur le résultant. Il faudrait déterminer par des preuves théoriques et des comparaisons expérimentales à partir de quelle valeur de n notre algorithme est plus performant. Nous conjecturons qu'il en est ainsi dès $n = 2$.

(2) Les matrices Φ et L ressemblent à des matrices de Sylvester. Plus précisément, la matrice de Sylvester est la matrice Φ dans le cas où $n = k = 2$. Or on sait effecteur une réduction de Gauss sur une matrice de Sylvester en un nombre d'opérations de l'ordre du carré de sa dimension. Comment utiliser, dans le cas général, la structure des matrices Φ et L pour accélérer la réduction?

9. Conclusion

Pour terminer, nous voudrions revenir sur les avantages et les limites de la méthode qu nous avons décrite.

L'inconvénient principal est certainement de ne s'appliquer que lorsqu'il n'y a qu'un nombre fini de solutions. Cela peut être gênant lorsque cette non-finitude provient uniquement des solutions à l'infini qui peuvent être sans intérêt pour le problème considéré. La généralisation ultérieure de la méthode au cas d'une infinité de solutions semble peu probable; elle nécessiterait, en tout cas, tout l'arsenal de la géométrie algébrique moderne.

Nous avons déjà mentionné que notre méthode permettait d'obtenir tous les renseignements algébriques et géométriques possibles sur l'ensemble des solutions, mais il nous semble important de souligner que l'algorithme lui même est 'géométrique'. Précisons ce qu'il faut entendre par là: tout problème raisonnable sur les polynômes, par exemple la résolution des systèmes d'équations algébriques, est, en un certain sens, invariant par changement linéaire de variables. C'est cette invariance qui constitue la nature géométrique du problème. Un algorithme peut donc être qualifié de géométrique s'il n'occulte pas cette invariance. Il n'est peut être

pas inutile de remarquer que celle-ci, qui provient ici de la linéarité en les U_i, est non seulement conservée, mais est utilisée d'une manière essentielle au paragraphe 6. Il nous semble que la nature géométrique des problèmes sur les polynômes a été insuffisamment utilisée dans la plupart des algorithmes les concernant, et que, dans la plupart des cas, des méthodes 'géométriques' devraient être beaucoup plus efficientes que les méthodes classiques où les polynômes sont considérés comme des polynômes à une variable à coefficients polynômes.

A part la première étape, la complexité de notre méthode dépend du nombre exact de solutions. Cet aspect ne doit pas être négligé pour une étude complète de complexité, car un problème intéressant n'a en général que peu de solutions.

Pour améliorer notre méthode, et calculer avec précision sa complexité, il faudrait étudier en priorité les points suivants:

- Comment utiliser la structure particulière de la matrice Φ, notamment lorsque certains des polynômes donnés sont sporadiques, pour accélérer la première étape?
- Comment conduire la récursivité de l'étape du paragraphe 5, ou comment en calculer la complexité pour résoudre le paradoxe suivant? L'existence de solutions à coordonnées nulles semble augmenter la complexité, alors qu'en pratique celà simplifie la résolution.
- Comparaison avec les algorithmes existant. Nous n'avons pu effectuer cette comparaison faute de disposer de ceux-ci. C'est un aspect non négligeable de cette comparaison que de constater qu'il est facile d'écrire un code pour notre algorithme en utilisant les variables flottantes de n'importe quel langage de programmation, alors que les algorithmes antérieurs nécessitent au préalable d'utiliser ou de mettre au point des programmes de maniements de polynômes.

Appendice A.1. Démonstration du théorème 3.1

Le théorème 3.1 se déduit facilement des résultats que l'on trouve dans tous les traités d'algèbre commutative. Nous montrons ici comment ils se déduisent de celui de Kaplansky [5].

Proposition A.1. *Avec les notations du paragraphe* 3, *si L est une extension algébriquement close de K, l'application* $(b_1, \ldots, b_n) \mapsto (X_1 - b_1, \ldots, X_n - b_n)$ *définit une bijection de l'ensemble des solutions dans L du système* (8) *sur l'ensemble des idéaux maximaux de* B_L.

C'est le théorème des zéros de Hilbert [5, théorème 32].

Proposition A.2. *Les assertions* (i) *et* (iii) *du théorème* 3.1. *sont équivalentes.*

Si $\dim_K B = \dim_{\bar{K}}(B_{\bar{K}})$ est fini, la longueur de $B_{\bar{K}}$ est finie et tout idéal premier de $\bar{B}$ est maximal et minimal [5, théorème 89]. Ces idéaux sont donc en nombre fini [5,

théorème 88], et il en est de même des solutions dans $\bar{K}$ du système (8) (proposition A.1). Réciproquement, si $B_{\bar{K}}$ n'a qu'un nombre fini d'idéaux maximaux, il est de dimension zéro [5, § 1–3, exemple 4] car c'est an anneau de Hilbert obtenu comme quotient d'un anneau de polynômes [5, théorème 31]; cet anneau $B_{\bar{K}}$ est donc de longueur finie [5, théorème 89], et donc aussi de dimension finie en tant que $\bar{K}$ espace vectoriel: un $B_{\bar{K}}$-module simple est isomorphe à un quotient de $B_{\bar{K}}$ par un idéal maximal; un tel quotient est isomorphe à $\bar{K}$ (proposition A.1), et est donc un $\bar{K}$-espace vectoriel de dimension 1 Nous avons ainsi montré que les quotients successifs d'une suite de composition de $B_{\bar{K}}$ sont des $\bar{K}$-espaces vectoriels de dimension finie; ceci entraîne immédiatement qu'il en est de même de $B_{\bar{K}}$.

Il est maintenant facile de terminer la démonstration du théorème 3.1: il est clair que (ii) entraîne (i); si (iii) est vérifiée, on a

$$\dim_L(B_L) = \dim_K(B),$$

et la proposition A.2 appliquée à l'anneau B_L montre que le système (8) n'a qu'un nombre fini de solutions dans une clôture algébrique de L, et donc aussi dans L.

Soit $(\xi_1, \ldots, \xi_n)$ une solution dans un corps L; quitte à agrandir L, on peut supposer que L est une extension algébriquement close de $\bar{K}$. Considérons l'idéal maximal $\mathfrak{m} = (X_1 - \xi_1, \ldots, X_n - \xi_n)$ de B_L (proposition A.1); l'intersection $\mathfrak{m} \cap B_{\bar{K}}$ est un idéal premier de $B_{\bar{K}}$ qui est maximal si (iii) est vérifié (démonstration de la proposition A.2), et donc de la forme $(X_1 - \zeta, \ldots, X_n - \zeta_n)$ avec les ζ_i dans $\bar{K}$ (proposition A.1). Il est facile de vérifier que l'on a $\xi_i = \zeta_i$ pour tout i, ce qui montre que les ξ_i sont algébriques sur K.

Les inégalités sur les nombres de solutions étant maintenant évidentes, il reste à montrer que le nombre de solutions sur $\bar{K}$ est majoré par $\dim_K(B) = \dim_{\bar{K}}(B_{\bar{K}})$. En montrant que $\dim_{\bar{K}}(B_{\bar{K}})$ était fini, on a, en fait, montré que ce nombre était la longueur de $B_{\bar{K}}$. Si $\mathfrak{m}$ est un idéal maximal de $B_{\bar{K}}$, le module simple $B_{\bar{K}}/\mathfrak{m}$ apparaît nécessairement comme quotient dans une suite de composition de $B_{\bar{K}}$, et donc dans toute; ceci montre que le nombre d'idéaux maximaux (i.e. de solutions) est majoré par la longueur de $B_{\bar{K}}$ (i.e. $\dim_K(B)$), ce qui termine la démonstration du théorème 3.1.

Appendice A.2. Démonstration du théorème 3.2.

Pour démontrer ce théorème, nous avons besoin d'une autre forme du théorème des zéros.

Proposition A.3. *Pour toute extension L de K, l'application qui, à $(a_0, \ldots, a_n)$, associe l'idéal engendré par les $a_iX_j - a_jX_i$ tels que $0 \leqslant i, j \leqslant n$ est une application injective de l'ensemble des solutions (projectives) dans L du système (9) dans*

l'ensemble des idéaux premiers gradués de A_L, maximaux parmi ceux qui ne contiennent pas A_L^1.

Si L est algébriquement clos, cette application est une bijection.

Si $(a_0, \dots, a_n)$ est une solution, il existe un i tel que $a_i \neq 0$; supposons, par exemple, que $i = 0$, et appelons I l'idéal engendré par les $a_iX_j - a_jX_i$ dans $L[X_0, \dots, X_n]$. Cet idéal est engendré par les

$$X_j - \frac{a_j}{a_0}X_0$$

car

$$a_iX_j - a_jX_i = a_i\left(X_j - \frac{a_j}{a_0}X_0\right) - a_j\left(X_i - \frac{a_i}{a_0}X_0\right).$$

Il en résulte que $L[X_0, \dots, X_n]/I$ est isomorphe à $L[X_0]$, ce qui montre que I est premier et que le seul idéal premier gradué contenant I est celui qui est engendré par les X_i.

Comme les F_i sont homogènes, l'égalité

$$F_i(a_0, \dots, a_n) = 0$$

implique

$$F_i\left(a_0\frac{X_0}{a_0}, a_1\frac{X_0}{a_0}, \dots, a_n\frac{X_0}{a_0}\right) = 0.$$

Autrement dit $F_i \equiv 0 \bmod (X_0 - a_0(X_0/a_0), \dots, X_n - a_n(X_0/a_0))$, ce qui montre que I contient tous les F_i, et que, modulo les F_i, l'idéal I est bien un idéal gradué de A_L premier et maximal parmi les idéaux premiers gradués ne contenant pas A_L^1.

Soit I' est l'idéal engendré par les $a_i'X_j - a_j'X_i$ pour une autre solution $(a_0', \dots, a_n')$. Il est facile de voir que $I = I'$ si et seulement si les a_i' sont proportionnels aux a_i, ce qui montre l'injectivité de l'application.

Supposons L algébriquement clos, et considérons un idéal premier gradué J de A_L ne contenant pas A_L^1. Le théorème des zeros [5, théorème 33] montre qu'il existe $a_0, \dots, a_n$ dans L, non tous nuls qui annulent tous les éléments de J, ce qui entraîne que J est contenu dans l'idéal engendré par les $a_iX_j - a_jX_i$ et que l'application est surjective.

Proposition A.4. *Soient A un anneau gradué tel que A^0 soit un corps K, et $\mathfrak{m}$ l'ideal engendré par A^1:*

(a) *L'idéal $\mathfrak{m}$ est le seul idéal premier gradué si et seulement si il existe un entier D tel que $A^d = 0$ pour $d \geq D$.*

(b) *Les idéaux premiers gradués autre que $\mathfrak{m}$ sont tous minimaux si et seulement si il existe un entier D tel que $\dim_K(A^d) = \dim_K(A^D)$ pour tout $d \geq D$.*

Si on localise l'anneau A en l'idéal premier $\mathfrak{m} = AA_1$, il est immédiat que

$$\dim_K(A_{\mathfrak{m}}/\mathfrak{m}^n A_{\mathfrak{m}}) = \dim_K(A/\mathfrak{m}^n) = \sum_{i=0}^{n-1} \dim_K(A^i).$$

Ainsi, si A^d est nul pour d assez grand, l'anneau A est de dimension 0 et $\mathfrak{m}$ est un idéal premier minimal [10, chapitre III, théorème 1]. Si $\dim_K(A^d)$ est constant pour d assez grand, tout idéal premier strictement contenu dans $\mathfrak{m}$ est minimal (ibid). Réciproquement, comme tout idéal premier minimal de A est gradué [2, chapitre IV, § 3, proposition 1], si $\mathfrak{m}$ est le seul idéal premier gradué, c'est un idéal premier minimal et l'anneau $A_{\mathfrak{m}}$ est de dimension zéro. Pour démontrer la réciproque de (b), considérons un idéal premier minimal (gradué) $\mathfrak{p}$ de A, et un élément homogène x qui n'appartient pas à $\mathfrak{p}$. Un idéal premier minimal contenant $\mathfrak{p}$ et x est gradué [2, loc. cit.] et donc égal à $\mathfrak{m}$ si tous les autres idéaux premiers gradués sont minimaux. Il en résulte qu'il n'existe aucun idéal premier entre $\mathfrak{p}$ et $\mathfrak{m}$ [5, théorème 142] et on conclut par le théorème cité de [10].

Nous pouvons maintenant démontrer le théorème 3.2.

(iv)⇒(iii): L'assertion (iv) entraîne que $\dim_L(A_L^d)$ décroît pour $d \geq D'$; il existe donc un entier D tel que $\dim_L(A_L^d) = \dim_L(A_L^D)$ pour tout $d \geq D$. L'assertion (iii) en résulte car $\dim_K(A^d) = \dim_L(A_L^d)$ pour tout d.

(iii)⇒(ii): Si (iii) est vérifié, on a $\dim_L(A_L^d) = \dim_K(A^D)$ pour toute extension L de K et tout entier $d \geq D$. Si $(a_0, \ldots, a_n)$ est une solution dans L du système (9), la proposition A.4 entraîne donc que l'idéal premier engendré par les $a_iX_j - a_jX_i$ est minimal. Ces idéaux sont donc en nombre fini [5, théorème 38], ce qui entraîne l'assertion (ii) (proposition A.3).

(ii)⇒(i): Evident.

(i)⇒(iv): Montrons d'abord l'assertion (iv) en nous limitant aux corps infinis L contenus dans $\bar{K}$. Si $(a_0, \ldots, a_n)$ est une solution dans $\bar{K}$, le sous-espace vectoriel de $A_{\bar{K}}^1$ engendré par les $a_iX_j - a_jX_i$ est un hyperplan dont l'intersection avec A_L^1 est un sous-espace vectoriel propre. Comme il n'y a qu'un nombre fini de tels sous-espaces vectoriels, il existe $y = \sum y_iX_i$ dans A_L^1 qui n'appartient à aucun d'eaux.

Comme y est homogène, les idéaux premiers de $A_{\bar{K}}$, minimaux parmi ceux contenant y sont gradués; comme $\mathfrak{m}A_{\bar{K}}$ est le seul idéal premier gradué contenant y (proposition A.3), il existe un entier d tel que $\mathfrak{m}^d A_{\bar{K}} \subset A_{\bar{K}}y$. Par ailleurs l'annulateur de y dans $A_{\bar{K}}$ est engendré par un nombre fini d'éléments homogènes $z_1, \ldots, z_n$; soit d' le plus grand degré des z_i. Ainsi, si z est un élément homogène de $ann_{\bar{K}}(y)$ de degré $\geq d + d'$, on a:

$$z \in \sum_i \mathfrak{m}^d A_{\bar{K}} z_i \subset \sum_i A_{\bar{K}} y z_i = 0.$$

Considérons d'autre part l'anneau $B = A_L/(y-1)A_L$; il satisfait aux conditions du théorème 3.1, car il y a bijection entre les solutions projectives du système (9) et les solutions du système obtenu en adjoignant l'équation $\sum y_iX_i = 1$ au système (9). Ainsi, B est un L-espace vectoriel de type fini et il existe un entier d'' tel que tout

élément de B soit image d'un élément de A_L de degré $\leqslant d''$. Posons $D' = \max(d'', d+d')$. Si $t \in A_L^i$ pour $i > D'$, on a donc

$$t = (y-1)u + v$$

avec degré$(v) \leqslant d'' <$ degré(t). Si u' est la partie homogène de plus haut degré de u, on a $t = yu'$, car si $yu' = 0$, on aurait degré $(u') \geqslant d + d'$ et donc $u' = 0$.

L'assertion (iv) est donc démontrée lorsque $L \subset \bar{K}$, et on a donc (i)$\Rightarrow$(iii). Si L n'est pas contenu dans $\bar{K}$ et si (i) est vérifié, $\dim_L(A_L^d)$ est donc constant pour d assez grand; on peut appliquer l'équivalence déjà démontrée en substituant L à K, ce qui démontre complètement (iv).

(a): Résulte immédiatement de (iii) et (iv).

(b) Reprenons les notations de la démonstration de (i)$\Rightarrow$(iv); les solutions projectives du système (8) sont en bijection avec les solutions du système obtenu en adjoignant $y = 1$ au système (9); leur nombre est donc majoré par $\dim_L(B)$. Montrons que, si $d > D'$ la surjection de A_L sur B induit une bijection de A_L^d sur B: soit $z \in A_L^d$; si on avait $z = (y-1)t$, le terme de plus haut degré de t serait annulé par y (car z est homogène); ceci montre que l'application $A_L^d \to B$ est injective. Tout élément de B est somme d'éléments qui sont images d'éléments homogènes de A_L; or, si $z \in A_L^i$ avec $i < d$, on a $z \equiv y^{d-i}z \bmod (y-1)$; si $z \in A_L^i$ avec $i > d$, on a, d'après (iv), $z = y^{i-d}t$, ce qui montre que l'application $A_L^d \to B$ est une bijection et que $\dim_K(A^d) = \dim_L(B)$.

(c) Toute solution dans L du système (9) est proportionnelle à une solution qui vérifie $y = 1$. D'après le théorème 3.1, cette solution est algébrique sur K, ce qui démontre (c).

(d) Immédiat.

(e) C'est la proposition A.4(a).

Appendice A.3. Démonstration du théorème 3.3

Le théorème 3.3 résulte facilement du théorème 1 de [7]; il n'est peut-être pas inutile d'en donner une deémonstration directe.

En raison du théorème 3.2, on peut supposer le corps K infini. La démonstration se fait par récurrence sur le nombre n d'indéterminées. Pour cela, posons

$$R = K[X_0, \ldots, X_n],$$

et appelons I l'idéal de R engendré par $F_0, \ldots, F_{k-1}$. L'anneau A que nous étudions est donc R/I. Considérons le complexe de Koszul ou de l'algèbre extérieure de I (pour la terminologie et les notions de base, on peut consulter n'importe quel livre d'algèbre homologique, par exemple [9]); c'est le complexe

$$\Lambda : 0 \to \Lambda_k \xrightarrow{\delta_k} \Lambda_{k-1} \xrightarrow{\delta_{k-1}} \cdots \xrightarrow{\delta_2} \Lambda_1 \xrightarrow{\delta_1} \Lambda_0 \to 0$$

dans lequel Λ_α est un R-module libre ayant pour base les éléments $e_{i_1} \wedge \cdots \wedge e_{i_\alpha}$ tels que $0 \leqslant i_1 < i_2 < \cdots < i_\alpha < k$, et $\Lambda_0 = R$; l'application δ_α est définie par

$$\delta_\alpha(e_{i_1} \wedge \cdots \wedge e_{i_\alpha}) = \sum_{j=1}^{\alpha} (-1)^{j+1} F_j e_{i_1} \wedge \cdots \wedge e_{i_{j-1}} \wedge e_{i_{j+1}} \wedge \cdots \wedge e_{i_\alpha}$$

et

$$\delta_1(e_i) = F_i$$

Si on affecte le degré $d_{i_1} + \cdots + d_{i_\alpha}$ à $e_{i_1} \wedge \cdots \wedge e_{i_\alpha}$ (d_i est le degré de F_i), les modules Λ_i sont gradués et les applications δ_i sont homogènes; ainsi, pour tout degré d, les restrictions δ_i^d des δ_i aux parties homogènes Λ_i^d, de degré d, des Λ_i forment un complexe Λ^d de K-espaces vectoriels de type fini. Notons $H_i^d = \ker(\delta_i^d)/\mathrm{im}(\delta_{i+1}^d)$ ses modules d'homologie; on a $H_0^d = A^d$.

Supposons les conditions équivalentes du théorème 3.2 vérifiées, et choisissons y dans A^1, tel que la multiplication par y de A^d dans A^{d+1} soit bijective pour d assez grand (rappelons que l'on *a* supposé K infini); cet élément y provient d'un élément Y de R^1, auquel on peut associer la suite exacte

$$0 \to R^{d-1} \xrightarrow{Y} R^d \to (R/YR)^d \to 0.$$

En tensorisant le complexe Λ^d par cette suite exacte, on obtient une suite exacte de complexes, et donc une suite exacte d'homologie

$$\cdots \to \bar{H}_{i+1}^d \to H_i^{d-1} \xrightarrow{Y} H_i^d \to \bar{H}_i^d \to \cdots$$

où $\bar{H}_i^d$ désigne l'homologie du complexe $\bar{\Lambda} = \Lambda \otimes (R/YR)$. Mais $\bar{\Lambda}$ est le complexe de Koszul de l'idéal $\bar{I}$ engendré par les images de $F_0, \ldots, F_{k-1}$ dans l'anneau R/YR qui est isomorphe à $K[X_0, \ldots, X_{n-1}]$. Aussi le théorème 3.3. va se déduire du théorème A.5 ci-dessous qui va lui-même se démontrer par récurrence sur n.

Théorème A.5. *Supposons les F_i rangés par degrés décroissants et non tous nuls. Si $A^d = 0$ pour tout d suffisamment grand, alors $H_i^d = 0$,*

(a) *pour tout d si $i \geqslant k - n$,*

(b) *pour tout $d \geqslant d_0 + d_1 + \cdots + d_{i+n} - n$ si $i < k - n$.*

Supposons ce résultat démontré pour $n = 0$ et montrons-le par récurrence, pour tout n. L'hypothèse de récurrence entraîne que $\bar{H}_{i+1}^d = 0$ pour tout d si $i + 1 \geqslant k - n + 1$ et pour $d \geqslant d_0 + \cdots + d_{i+n} - n + 1$ si $i < k - n$. Supposons que $H_i^d = 0$ pour les grandes valeurs de d. La suite exacte

$$\bar{H}_{i+1}^d \to H_i^{d-1} \to H_i^d$$

montre, par récurrence descendante sur d que $H_i^{d-1} = 0$ pour tout d si $i \geqslant k - n$ et pour $d \geqslant d_0 + \cdots + d_{i+n} - n + 1$ si $i < k - n$. On est donc ramené aux lemmes suivants:

Lemme A.6. *Si $A^d = 0$ pour d assez grand, il en est de même de H_i^d, pour tout i.*

L'hypothèse entraîne que l'idéal $\mathfrak{m} = (X_0, \ldots, X_n)$ est le seul idéal premier contenant I. Si $\mathfrak{p}$ est un autre idéal premier de R, le complexe $\Lambda \otimes R_\mathfrak{p}$ est le complexe de Koszul de l'idéal $I \otimes R_\mathfrak{p}$ du localisé $R_\mathfrak{p}$; mais $I \otimes R_\mathfrak{p} = R_\mathfrak{p}$ et $\delta_1 \otimes R_\mathfrak{p}$ est surjectif. Soit ε dans $\Lambda_1 \otimes R_\mathfrak{p}$ tel que $(\delta_1 \otimes R_\mathfrak{p})(\varepsilon) = 1$. Posant $\varepsilon_\alpha(e_{i_1} \wedge \cdots \wedge e_{i_\alpha}) = \varepsilon \wedge e_{i_1} \wedge \cdots \wedge e_{i_\alpha}$, on vérifie facilement que $\varepsilon_{\alpha-1} \circ (\delta_\alpha \otimes R_\mathfrak{p}) + (\delta_{\alpha-1} \otimes R_\mathfrak{p}) \circ \varepsilon_\alpha = id(\Lambda_\alpha \otimes R_\mathfrak{p})$, ce qui montre que l'homologie $H_i \otimes R_\mathfrak{p}$ de $\Lambda \otimes R_\mathfrak{p}$ est nulle pour tout $\mathfrak{p} \neq \mathfrak{m}$. Ainsi, H_i est annulé par une puissance de $\mathfrak{m}$, ce qui entraîne le résultat.

Lemme A.7. *Le théorème A.5 est vrai pour $n = 0$.*

L'assertion (a) est immédiate. Si $i < k$, le lemme précédent montre que, si $x \in \ker \delta_i$, il existe un entier h tel que $X_0^h x = \delta_{i+1}(z)$. Les éléments de base de Λ_{i+1} étant au plus de degré $d_0 + d_1 + \cdots + d_i$, leurs coefficients sont divisibles par X_0^h si degré$(x) \geq d_0 + \cdots + d_i$. On peut alors diviser par X_0^h, ce qui montre que $x \in \mathrm{im}(\delta_{i+1})$.

Le théorème 3.3 se déduit maintenant très facilement du théorème A.5 et de la suite exacte

$$\bar{H}_1^d \to H_0^{d-1} \xrightarrow{Y} H_0^d \to \bar{H}_0^d.$$

En effet, le théorème 3.2 entraîne que $(R/YR)^d$ est nul pour d assez grand; il en résulte par le théorème A.5 que $\bar{H}_i^d = 0$ pour tout d si $i \geq k - n + 1$ et pour d $d \geq d_0 + d_1 + \cdots + d_{i+n-1} - n + 1$ sinon. Ceci entraîne que la multiplication par Y est surjective si $d \geq d_0 + \cdots + d_{n-1} + 1 - n$ et injective si $k = n$ ou si $d - 1 \geq d_0 + \cdots + d_n - n$. Ainsi, on peut prendre (notations du théorème 3.2)

$$D' = d_0 + d_1 + \cdots + d_{n-1} + 1 - n$$

et

$$D = d_0 + d_1 + \cdots + d_{n-1} - n \quad \text{si } k = n$$

et

$$D = d_0 + d_1 + \cdots + d_n - n \qquad \text{si } k > n$$

(les hypothèses entraînent que $k \geq n$).

Avec la convention que $d_n = 1$ si $k = n$, on a donc $\max(D, D') = d_0 + \cdots + d_n - n$, ce qui démontre le théorème 3.3.

Appendice A.4

Nous donnons ici une forme explicite de l'algorithme des paragraphes 4 et 5. Les commentaires entre rochets permettent de faire le lien entre les deux présentations de l'algorithme et utilisent les notations des paragraphes 4 et 5.

1. Données d'entrée

Ce sont:
- le nombre K des polynômes,
- le nombre N des variables,
- la liste $D(0), D(1), \ldots, D(K-1)$ des degrés des polynômes,
- les polynômes eux-mêmes, codés par exemple par la liste de leurs coefficients dans l'ordre lexicographique, soit $F(I, 1), \ldots, F(I, M)$ pour le I-ème polynôme, avec

$$M = \binom{D(I)+N}{N}.$$

2. Initialisation

(I1) Si $K < N$ alors faire; écrire 'il y a trop de variables'; fin du programme; fait;
(I2) $D(K) \leftarrow 1$;
(I3) Soit $E(0), \ldots, E(K)$ les $D(I)$ rangés dans l'ordre décroissant;
(I4) $\mathrm{DD} \leftarrow E(0)+E(1)+\cdots+E(N)-N$;
(I5) $\mathrm{NL} \leftarrow (\mathrm{DD}+1)*(\mathrm{DD}+2)*\cdots*(\mathrm{DD}+N)/(1*2*\cdots*N)$;
[NL est le nombre de ligne de la matrice ϕ = dimension R^D]
(I6) Créer C, une matrice NL×NL initialisée en matrice unité;
[ce sera la matrice C du paragraphe 4]
(I7) Créer un vecteur PHI de dimension NL;
[ce sera la colonne courante de ϕ]

3. Réduction de la matrice Φ

(R1) $B(I) \leftarrow 0$ pour $I = 1$ à NL; [les I pour lesquels $B(I) \neq 0$ seront les numéros des lignes de Φ où on a déjà trouvé un pivot]
(R2) Faire pour $I = 0$ à $K-1$;
(R3) $\mathrm{DI} \leftarrow \mathrm{DD}-D(I)$; $\mathrm{NC} \leftarrow (\mathrm{DI}+1)*(\mathrm{DI}+2)*\cdots*(\mathrm{DI}+N)/(1*2*\cdots*N)$;
[NC est le nombre de colonnes de Φ correspondant au I-ème polynôme]
(R4) Faire pour $J = 1$ à NC;
(R5) $\mathrm{PHI}(J1) \leftarrow 0$ pour $J1 = 1$ à NL; [réinitialisation de PHI]
(R6) Soit $E(1), \ldots, E(N)$ les exposants du J-ème monôme de degré DI (pour l'ordre lexicographique);
(R7) Faire pour $J1 = 1$ à NL;
(R8) Soit $\mathrm{EE}(1), \ldots, \mathrm{EE}(N)$ les exposants du $J1$-ème monôme de degré DD;
(R9) Si, pour $J2 = 1$ à N on a $\mathrm{EE}(J2) \geqslant E(J2)$, alors $\mathrm{PHI}(J1) \leftarrow F(I, J3)$) où $F(I, J3)$ est le coefficient du monôme d'exposants $\mathrm{EE}(1)-E(1), \ldots, \mathrm{EE}(N)-E(N)$ dans le I-ème polynôme;
(R10) Fait (R7); [PHI est maintenant la J-ème colonne de la partie de Φ correspondant au I-ème polynôme]
(R11) $\mathrm{PHI} \leftarrow C*\mathrm{PHI}$; [L'opération $*$ désigne le produit de matrices]

(R12) Si c'est possible, choisir $J1$ tel que $1 \leqslant J1 \leqslant$ NL et PHI$(J1) \neq 0$ et $B(J1) = 0$; [PHI$(J1)$ est le pivot; le choisir de valeur absolue maximale si les données sont réelles ou complexes; le choisir de numérateur et de dénominateur petits si les données sont rationnelles]
Sinon aller en (R15);

(R13) $B(J1) \leftarrow 1$;

(R14) Pour $J2 = 1$ à NL, si $B(J2) = 0$ et PHI$(J2) \neq 0$ alors pour $J3 = 1$ à NL, $C(J2, J3) \leftarrow C(J2, J3) - (\text{PHI}(J2)/PHI(J1)) * \text{PHI}(J1, J3)$; réduction de Gauss; si les données sont des entiers ou des polynômes, il y a lieu de multiplier les deux membres par PHI$(J1)$ et de diviser le résultat par le pivot précédent, afin de limiter la croissance de la taille des données]

(R15) Fait (R4);

(R16) Fait (R2); [la valeur actuelle de C est celle qui est considérée au paragraphe 4]

(R17) SR $\leftarrow$ card$\{I; B(I) = 0\}$; [SR est l'entier $s - r$ du paragraphe 4]

(R18) Si SR $= 0$ alors faire; écrire 'aucune solution'; fin du programme; fait;

(R19) Soit CC la matrice $R \times$ NL obtenue en supprimant de C les lignes telles que $B(J1) = 1$; [c'est la matrice des '$s - r$ dernières lignes de C']

4. Réduction de Λ

[Il faut d'abord écrire cette matrice Λ]

(C1) NC $\leftarrow$ NL $*$ DD/(DD $+ N$); [nombre de colonnes de Λ]

(C2) Créer LAM, une matrice SR $\times$ NC $\times (N + 1)$; [LAM sera considérée comme une matrice SR $\times$ NC de vecteurs de dimension $N + 1$, les composantes d'un tel vecteur étant les coefficients d'un polynôme de degré 1]

(C3) Faire pour $I = 1$ à NC;

(C4) Faire pour $J = 0$ à N;

(C5) Pour $I1 = 1$ à SR, LAM$(I1, I, J) =$ CC$(I1, I2)$ où $I2$ est le numéro parmi les monômes de degrés au plus DD du produit de la J-ème variable et du monôme de degré au plus DD $- 1$ de numéro I (par convention, la 0-ème variable est l'élément 1);

(C6) Fait (C4);

(C7) Fait (C3);

(C8) Appeler la procédure REDUC; [La réduction étant faite récursivement, la procédure REDUC s'appelle elle-même]

Procédure REDUC:

(F1) LPIV $\leftarrow 0$; [ligne du dernier pivot]

(F2) Pour $J = 0$ à N faire; [J est l'indice de la variable courante]

(F3) CPIV $\leftarrow 0$; [colonne du dernier pivot]

(F4) Si, pour $I1 =$ LPIV $+ 1$ à SR et pour $I =$ CPIV $+ 1$ à NC on a LAM$(I1, I, J) = 0$, alors aller en (F11); Sinon, choisir $I1$ et I tels que LAM$(I1, I, J) \neq 0$; [pivot]

(F5) CPIV ← CPIV + 1; LPIV ← LPIV + 1;

(F6) Permuter les lignes $I1$ et LPIV et les colonnes I et CPIV de LAM;

(F7) Pour $I2 = 1$ à SR, pour $I3 = \text{CPIV} + 1$ à NC, pour $I4 = 0$ à N, $\text{LAM}(I2, I3, I4) \leftarrow \text{LAM}(I2, I3, I4) - (\text{LAM}(\text{LPIV}, I3, J)/\text{LAM}(\text{LPIV}, \text{CPIV}, J)) * \text{LAM}(I2, \text{CPIV}, I4)$; [réduction de Gauss sur les colonnes]

(F8) Si LPIV = SR alors aller en (F13);

(F9) Pour $I2 = \text{LPIV} + 1$ à SR, pour $I3 = 1$ à NC pour $I4 = 0$ à N, $\text{LAM}(I2, I3, I4) \leftarrow \text{LAM}(I2, I3, I4) - (\text{LAM}(I2, \text{CPIV}, J)/\text{LAM}(\text{LPIV}, \text{CPIV}, J)) * \text{LAM}(\text{LPIV}, I3, I4)$; [réduction de Gauss sur les lignes]

(F10) Aller en (F4);

(F11) Fait (F2);

(F12) Si LPIV < SR alors faire; ecrire 'il y a une infinité de solutions'; fin du programme; fait; [en effet le rang de LAM est inférieur à son nombre de lignes (voir paragraphe 6.1(a)]

(F13) Vérifier que, pour $I1 = \text{SR} - \text{CPIV} + 1$ à SR, pour $I2 = \text{CPIV} + 1$ à NC, pour $I3 = 0$ à N, $\text{LAM}(I1, I2, I3) = 0$; sinon faire; ecrire 'il y a une infinité de solutions'; fin du programme; fait; [il s'agit de la vérification (b) du paragraphe 6.1]

(F14) Mémoriser la sous-matrice de LAM constituée des coefficients $\text{LAM}(I1, I2, I3)$ tels que $I1 > \text{SR} - \text{CPIV}$ et $I2 \leqslant \text{CPIV}$;

(F15) Supprimer les coefficients $\text{LAM}(I1, I2, I3)$ de LAM tels que $I1 > \text{SR} - \text{CPIV}$ ou $I2 \leqslant \text{CPIV}$;

(F16) SR ← SR − CPIV; NC ← NC − CPIV;

(F17) Si $\text{SR} \neq 0$, alors appeller REDUC;

(F18) Fin de REDUC; [Le polynôme G du théorème 4.1 est le produit des déterminants des matrices mémorisées en (F14)]

Bibliographie

[1] N. Bourbaki, *Algèbre* (Hermann, Paris, nouvelle ed., 1970) chapitre II, § 10, définition 1.

[2] N. Bourbaki, *Algèbre Commutative* (Hermann, Paris, 1962) chapitres III et IV.

[3] G.E. Collins, Factoring univariate polynomials in polynomial average time, in: *Symbolic and Algebraic Computation*, Lecture Notes in Computer Science **72** (Springer, Berlin, 1979) 317–329.

[4] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics **52** (Springer, New York, 1977).

[5] I. Kaplansky, *Commutative Rings* (Allyn and Bacon, Boston, 1970).

[6] D. Lazard, Algèbre linéaire sur $K[X_1, \ldots, X_n]$ et élimination, *Bull. Soc. Math. France* **105** (1977) 165–190.

[7] D. Lazard, Systems of algebraic equations, in: *Symbolic and Algebraic Computation*, Lecture Notes in Computer Science **72** (Springer, Berlin, 1979) 88–94.

[8] F.S. Macaulay, *The Algebraic Theory of Modular Systems*, Cambridge Tracts in Mathematics **19** (Cambridge Univ. Press, Cambridge, 1916) et (Stechert-Hafner, New York, 1964).

[9] D.G. Northcott, *An Introduction to Homological Algebra* (Cambridge University Press, Cambridge, 1960).

[10] J.P. Serre, *Algèbre locale, multiplicités*, Lecture Notes in Mathematics **11** (Springer, Berlin, 1965).

[11] B.L. van der Waerden, *Moderne Algebra II* (Springer, Berlin, ed. 1 à 3, 1936, 1950, 1955); English translation: (Ungar, New York, 1950).
[12] D.Y.Y. Yun, On algorithms for solving systems of polynomials equations, *ACM SIGSAM Bull.* **27** (1973) 19–25.
[13] D. Lazard, Problem #7 and systems of algebraic equations, *ACM SIGSAM Bull.* **54** (1980) 26–29.